COMMUNITIES AND COMPETITION IN RANDOM NETWORKS

by

JORDAN CHELLIG

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
The University of Birmingham
March 2023

Abstract

We consider two problems inspired by the social properties of large-scale random networks. Firstly, we consider evolutionary games on a population whose underlying topology of interactions is determined by a binomial random graph. Our focus is on 2-player symmetric games with 2 strategies played between the incident members of such a population. Players update their strategies synchronously. At each round, each player selects the strategy that is the best response to the current set of strategies its neighbours play. We show rapid convergence to unanimity for p in a range that depends on a certain characteristic of the payoff matrix. In the case where the matrix possesses a further strategic bias, we determine a sharp threshold on p, above which the largest connected component reaches unanimity with high probability. For p below this critical value, where this does not happen, we identify those substructures inside the largest component that remain discordant throughout the evolution of the system. We consider extensions of this system into three or more strategies and declare unanimity for two specific cases depending on entries in the payoff matrix.

Our final project considers graph modularity: a quantity that has been introduced in order to quantify how close a network is to an ideal modular network. In such an ideal network the nodes form small interconnected communities that are joined together with relatively few edges. In this thesis, we consider this quantity on a recent probabilistic model of complex networks introduced by Krioukov et al. [50].

This model views a complex network as an expression of hidden hierarchies, encapsulated by an underlying hyperbolic space. For certain parameters, this model was proved to have typical features that are observed in complex networks such as power-law degree distribution, bounded average degree, clustering coefficient that is asymptotically bounded away from zero, and ultra-small typical distances. We investigate its modularity and we show that, in this regime, it converges to 1 in probability.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank Nikolaos Fountoulakis for his guidance, supervision and many insightful conversations across the last 3 and a half years, along with being an excellent collaborator. I would also like to thank my co-supervisor Richard Mycroft, who has always given me excellent guidance, and without whom I would probably not be on the path that I am now. I would also like to thank my excellent co-authors, Calina and Fiona.

I would also like to extend massive thanks and love to my favourite group of gossips, the BHT: Ben, Andy F, Andy P, Joe, Barry and Harry. They have shown me nothing but support and have given me years of fond, yet slightly questionable, memories. I couldn't do it without you!

I would also like to thank my dungeon buddies, Jamie, Rob, Andrew, Ricky and Eoin for providing me with a wonderful selection of insightful conversations and slightly more chaotic memories.

Finally, I would like to express a massive amount of gratitude to both my parents; my brother Jamie; and my sister Karima, for their constant support and encouragement. Finally to my dear Greg, I thank you for your unwavering support, care and affection; alongside being made to hear about random graphs, more than any English teacher ever should! This thesis is for all of you.

CONTENTS

CHAPTER 1

INTRODUCTION

We are concerned with the analysis of random graph models as a means to understand phenomena concerning social interactions. We consider two major problems: The former problem concerns the study of *best response dynamics*, a game theoretic evolutionary process, combined with a local network topology given by the *binomial random graph.* The latter problem is concerned with computing the *modularity* score of the *hyperbolic random graph.* These random graphs are suggested to be promising models for representing complex networks, possessing many typical features we would expect of large-scale real-world networks.

Each section within the introduction features some brief expository and foundational material which supports the main results. We then conclude with a theorem statement. The chapters are structured to lead with a relevant review of the current literature; followed by a proof of the main result; and finally a set of concluding remarks and suggestions for further work.

The work found within chapter two is joint with Calina Durbac and Nikolaos Fountoulakis, and has been published within [22]. The work within chapter three is joint with Nikolaos Fountoulakis. While the work within chapter four is joint with Nikolaos Fountoulakis and Fiona Skerman and has been published within [23].

1.1 Best Response Dynamics on Random Graphs

Our first result concerns the analysis of a collection of agents competing in evolutionary games, with an interaction topology given by the binomial random graph. Our main result concerns the speed at which the agents playing these games settle into a unanimous strategy. In Sections 1.1.1 and 1.1.2 we detail the key components of the system, namely the underlying topology of interactions and the mode by which pairs of agents interact. The subsequent subsections detail the rules of the system, followed by an outline of the general behaviour in a variety of regimes, namely Theorems 1.1.1 to 1.1.4.

1.1.1 The Binomial Random Graph

A graph $G = (V, E)$ consists of a set of vertices, and a set of unordered pairs of vertices referred to as edges. We denote the set of vertices of G as $V(G)$, and the set of edges as $E(G)$. In this section of the thesis we are primarily focused on the *binomial random graph*, which was first studied by Erdős and Rényi [29], along with Edgard Gilbert independently [37]. For $n \in \mathbb{N}$, we denote $\mathcal{G}_n$ to be the set of all graphs on n vertices. We define the binomial random graph as follows: for $n \in \mathbb{N}$ and $p \in [0, 1]$, we denote $G(n, p)$ to be a probability distribution on $\mathcal{G}_n$. We have for a given graph $G \in \mathcal{G}_n$,

$$\mathbb{P}(G) = p^{|E(G)|}(1-p)^{\binom{n}{2}-|E(G)|},$$

in the $G(n, p)$ model. This distribution describes a process which starts on an empty n vertex graph and for every pair of vertices we connect them with probability p independently of all other pairs. As a brief aside, we also mention the alternative model of the binomial random graph. For $n, m \in \mathbb{N}$ with $0 \leq m \leq \binom{n}{2}$, we denote $\mathcal{G}'_{n,m}$ to be the set of all graphs on n vertices and m edges. We define $G(n, m)$ to be a probability distribution over $\mathcal{G}'_{n,m}$. We have for a given $G \in \mathcal{G}'_{n,m}$,

$$\mathbb{P}(G) = \binom{\binom{n}{2}}{m}^{-1},$$

in the $G(n,m)$ model. This distribution describes a process starting with an initial empty graph on n vertices; sequentially m edges are added to the graph randomly (and uniformly) across the current non-edges. We remark that these two models are asymptotically equivalent in n, this can be seen by considering the $G(n,p)$ model with $p = m/\binom{n}{2}$, for more on $G(n,m)$ see [15]. Throughout this thesis, we will only consider the $G(n,p)$ model.

We briefly detail some well-known results regarding the $G(n,p)$ model, namely the phase transitions that occur for ranges of p. For different choices of $p_n = p(n)$, we have that $G(n,p_n)$ tends to have different graph properties with high probability. Initially, if $p_n = o(n^{-2})$ then $G(n,p)$ is empty with high probability, while for $p_n = \omega(n^{-2})$ we have that $G(n,p)$ contains at least one edge with high probability. The next threshold occurs at $p_n = 1/n$, known at the *critical threshold*. For $\varepsilon > 0$, we have that if $p_n = (1-\varepsilon)/n$ then, with high probability, all connected components of $G(n,p_n)$ have size $O(\log n)$, where $\log \cdot$ is the natural logarithm. While for $p_n = (1+\varepsilon)/n$ we have the emergence of a giant connected component, which is linear in size. The remaining components have order $O(\log n)$, see [15, 45, 52]. Figure 1.1 shows this transition: *In these Java simulations, a small radius indicates a higher relative degree.*

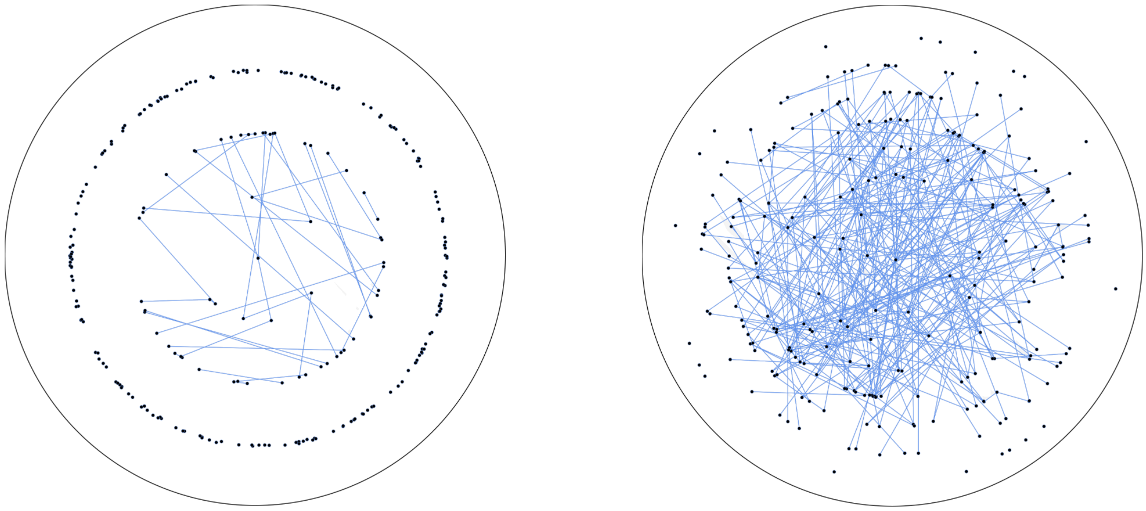

Figure 1.1: Two realisations of $G(n,p)$ with $n = 250$. We have $p = 0.0005$ on the left, and $p = 0.045$ on the right. This illustrates the sub-critical and super-critical regimes respectively.

Following the above, is the sharp connectivity threshold, due to Gilbert [37]. Let $\omega(n) : \mathbb{N} \to \mathbb{R}^+$, such that $\omega(n) \to \infty$ as $n \to \infty$. If $np_n = \log n + \omega(n)$, then with high probability $G(n, p_n)$ is connected. While if $np_n = \log n - \omega(n)$, then $G(n,p)$ almost surely possesses isolated vertices, this can be seen in Figure 1.2. We remark the threshold for a Hamilton cycle occurs slightly above the connectivity threshold at $np_n = \log n + \log\log n$ [51, 79].

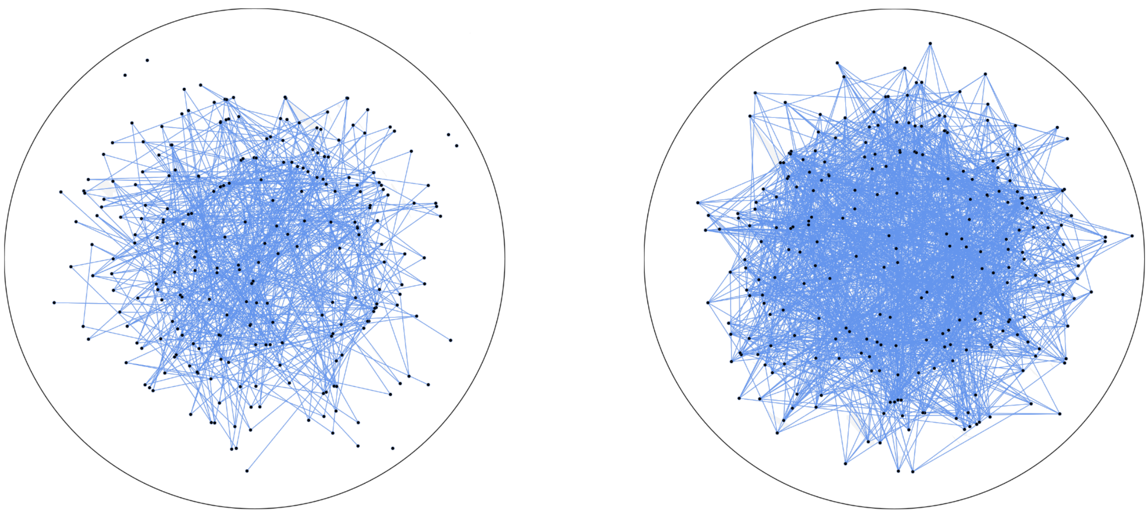

Figure 1.2: Two realisations of $G(n,p)$ with $n = 250$. We have $p = 0.01$ on the left, and $p = 0.0221$ on the right. The former possesses isolated vertices, while the latter does not.

Throughout this thesis, we will be focused on p ranging from just above the critical threshold, ($p_n = \omega(1)/n$). We will also consider the range where $p_n \geq \Lambda n^{-1/2}$, for a suitably large positive constant Λ.

1.1.2 Normal-Form Games

Normal-Form games are a class of mathematical games focusing on simultaneous decision-making. The general setup we consider is that of two players, each of whom can choose either strategy zero or one. Each player will then simultaneously choose a strategy, and depending on which of the four combinations arises each player will receive a given payoff. Each of the player's respective payoffs can be represented by a pair of real-valued 2×2 matrices, each referred to as the *payoff matrix*. Suppose we have two players, Alice and Bob. If Alice is assigned a payoff matrix A and plays strategy i, while Bob is assigned payoff matrix B and plays strategy j; then Alice receives payoff $a_{i,j}$, while Bob receives payoff $b_{i,j}$. The pair (A, B) is referred to as a *bi-matrix game* [68].

A prominent class of bi-matrix is the case where $A = -B$. In this case, the sum of Alice and Bob's payoff is always zero, known as a *zero-sum game*. A fundamental question we can ask is, given Alice and Bob have knowledge of (A, B) what would be their optimal choice of strategies, and what would be their respective payoffs? As Alice is choosing the row, it would be intuitive to expect her to choose the row which contains the largest element. However, Bob is choosing the columns and will aim to choose the smallest element in A, which corresponds to a large payoff in B. Thus, the only elements which can satisfy both players (in the sense that neither would have the incentive to switch strategy from the current logical position), is an entry which is both a row minimum and a column maximum, also referred to as a *saddle point*. In terms of pure strategies, where the game is only played once, the zero-sum game has a solution if and only if the payoff matrix has a saddle point [17]. We observe that this saddle point is also an example of a *Nash equilibrium*, wherein the strategy chosen by Alice is the optimal response to any optimal strategy chosen by Bob, and vice-versa. When a saddle point does not exist, we may instead consider mixed strategies, where Alice and Bob play the game repeatedly and their strategies are instead given as a probability distribution over their own pure strategies. In this case, determining the optimal distribution is a standard problem solved through Linear Programming, for more on general matrix games see, [68, 74].

Throughout this thesis, we will be focusing on symmetric matrix games. In this case, if Q is a 2×2 payoff matrix, the *symmetric game* is defined as the bi-matrix game (Q, Q^T). Immediately we note that if all the column maximums occur in the same row, then for Q^T, all the row maximums will occur in the same column. This causes Alice and Bob to immediately agree on the same strategy, and thus the game is solved. This type of argument is an example of dominating strategies, we will review this argument rigorously in Section 2.1.2. In the remaining cases (where neither strategy dominates), there are two Nash equilibria that will occur, either at the top left and bottom right corners, or bottom left and top right corners. The former type of matrix is known as a *co-ordination game*, Alice and Bob will receive the best payoff for choosing the same strategies; while in the latter, known as an *anti-coordination game*, they should choose different strategies. Generally, this analysis describes the broad strategy the pair should follow; however, it is not entirely clear how the two should decide between the Nash equilibria. In this thesis, we will be considering a type of iterative sequence of pure strategies, known as *best response*. In this setting, agents will play repeated games and will choose their next strategy by playing what should have been the best response to their opponent's current strategy. We can view these tactics through the perspective that both players believe that their opponent will utilise the same strategy in the following round, and thus they should react accordingly.

1.1.3 A Model for Best Response Dynamics on Graphs

Let $Q = [q_{i,j}]_{0 \leq i,j \leq 1}$ be a 2×2 payoff matrix with entries in $\mathbb{R}$. The rows and columns of Q are indexed by $\{0, 1\}$, which are assumed to represent the strategies. Each player will now choose a strategy from $\{0, 1\}$. Player 1 will then receive a payoff given by $q_{i,j}$, where i

is the strategy chosen by Player 1, and j is the strategy chosen by Player 2. Analogously, Player 2 will receive a payoff $q_{j,i}$.

The interaction topology of the agents/players is represented by a fixed graph $G = (V, E)$. The nodes of G represent the agents, and if two nodes are adjacent, then the corresponding agents interact with each other by playing the game with payoff matrix Q. We refer to this process as an *interacting node system* $(G, Q, \mathcal{S})$, which we detail formally as follows. We fix a graph G, a payoff matrix Q, and for every $v \in V(G)$ an initial vertex strategy $\mathcal{S}(v)$ where $\mathcal{S} : V(G) \to \{0, 1\}$. We consider a discrete time process. For each $t \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, we denote by $S_t(v)$ the strategy played by vertex v at step t. Thus, $S_t : V(G) \to \{0, 1\}$. Let $N_G(v)$ denote the set of neighbours of a vertex $v \in V(G)$ in G. For a vertex $v \in V(G)$, a step $t \in \mathbb{N}_0$ and $j \in \{0, 1\}$, we set $n_t(v; j) = |N_G(v) \cap \{u : S_t(u) = j\}|$. To each vertex $v \in V(G)$, we assign the initial state $S_0(v) = \mathcal{S}(v)$. To progress from round t to round $t + 1$, the following evolution rule is applied: Each vertex v will play a game against each one of its neighbours, playing strategy $S_t(v)$. For each game played, the vertex will receive a payoff given by the corresponding entry in the payoff matrix Q. We denote the *total payoff* for a vertex v at time step t to be the sum of all payoffs that v received in that round. Therefore, for a vertex v with $S_t(v) = i$, we define the *total payoff* of v at time t as $T_t(v) := n_t(v; 0)q_{i,0} + n_t(v; 1)q_{i,1}$. We define the *alternative payoff* of v at time t, as the total payoff the vertex would have received, had they played the other available strategy. Hence, the alternative payoff for a vertex v with $S_t(v) = i$ is given as: $T'_t(v) := n_t(v; 0)q_{1-i,0} + n_t(v; 1)q_{1-i,1}$. The value of $S_{t+1}(v)$ is determined by comparing the values of the current payoff and the alternative payoff. If the alternative payoff is strictly greater than the total payoff, then the vertex will switch to strategy $1 - i$ in round $t + 1$; otherwise, it will continue with its current strategy. For a vertex v at time t, we can succinctly express the evolution rule as follows:

$$S_{t+1}(v) = \begin{cases} S_t(v) & \text{if } T_t(v) \geq T'_t(v); \\ 1 - S_t(v) & \text{if } T_t(v) < T'_t(v). \end{cases}$$

We wish to analyse the global evolution of the strategies as time elapses. We say that a vertex's strategy is *periodic* if there exists some $T, p \in \mathbb{N}$, such that $S_T(v) = S_{T+kp}$ for all $k \in \mathbb{N}$. We call the least possible p which satisfies this definition the *period*. We say that the evolution of a system is *unanimous*, if there exists some $T \in \mathbb{N}$ such that for all $t \geq T$ and all distinct pairs of vertices $u, v \in V(G)$, we have that $S_t(u) = S_t(v)$. We say that the system is *stable* if there exists $T \in \mathbb{N}$ such that for all $v \in V(G)$ and $t \geq T$, we have that $S_T(v) = S_t(v)$.

1.1.4 Classification of Payoff Matrices

The evolution of the node system is largely governed by the entries present in the payoff matrix Q. As a consequence of choosing specific entries for Q, it can be the case that one strategy is strictly more beneficial than the other. In this case, the evolution of the system is trivial, as all agents will unanimously maximise their payoff; we capture this behaviour with the notion of degeneracy. We say that a payoff matrix $Q = (q_{i,j})$ with $i, j \in \{0, 1\}$ is *non-degenerate*, if one of the following holds:

(i) We have that $q_{0,0} > q_{1,0}$ and $q_{0,1} < q_{1,1}$.

(ii) Or, $q_{0,0} < q_{1,0}$ and $q_{0,1} > q_{1,1}$.

Otherwise, we say that Q is *degenerate*. The behaviour of interacting node systems with degenerate payoff matrices will be discussed in Section 2.1.2. Furthermore if Q is such that (i) holds, then we say that Q is in the *majority regime*, while if (ii) holds then we say that Q is in the *minority regime*. The context and implications for this classification are discussed within Section 2.1.1. We discuss some examples of this classification below:

Examples

1. *The Hawk-Dove Game.* In this game, a population of animals consists of two types of individuals which are differentiated by the amount of aggression they display during

their interactions. There is the most aggressive type (hawk) and the least aggressive, or most cooperative, type (dove). When two of them interact over some fixed resource, the outcome depends on their types. If two hawks compete over the resource they get injured, because of the fighting, and one of them (at random) manages to grab the resource. Hence, if $R > 0$ is the gain from the resource, each is expected to gain $R/2$ out of the fight, but they pay a price $P > R$ for their injuries, whereby their overall gain is $(R - P)/2$. If a hawk interacts with a dove, then the hawk grabs the resource gaining R, whereas the dove walks away with nothing. Finally, if two doves interact, then they share the resource each gaining $R/2$. Thus, if $Q_{\mathrm{H-D}}$ denotes the payoff matrix of the Hawk-and-Dove game, this is

$$Q_{\mathrm{H-D}} = \begin{bmatrix} \frac{R-P}{2} & R \\ 0 & \frac{R}{2} \end{bmatrix}, \tag{1.1}$$

where the first row and column correspond to the hawk strategy, and the second row and column correspond to the dove strategy. We observe that as $R < P$, the column maxima occur on the anti-diagonal. Thus Q_{H-D} is a non-degenerate payoff matrix in the minority regime.

2. *The Prisoners Dilemma.* In this game, two individuals are arrested while committing a crime together, but they are put in different cells. The police do not have enough evidence to convict them, but they make an offer to each one separately. If one of them confesses (defects) but the other remains silent (cooperates), then the former is released but the other is sentenced to imprisonment of $P > 0$ years. If they both confess, they are sentenced to $R > 0$ years of imprisonment, for $R < P$. Finally, if both remain silent, they are both sentenced to S years imprisonment for some minor offence, as the police do not have enough evidence. In this case, $S < R$.

Thus, $0 < S < R < P$. The payoff matrix of the Prisoners dilemma game Q_{PD} is

$$Q_{\mathrm{PD}} = \begin{bmatrix} -R & 0 \\ -P & -S \end{bmatrix},$$

where the first row and column correspond to the Defect strategy and the second row and column correspond to the Cooperate strategy. (Here, we put the $-$ sign in front of these quantities, as imprisonment is thought of as a loss.) Hence, this is a case of a degenerate payoff matrix.

1.1.5 Main Results and Theorems for Best Response Dynamics with Two Strategies

We consider an interacting node system where the underlying graph G is both random and suitably dense. We let $G(n,p)$ be the *binomial random graph* on vertex set $V_n := [n] := \{1, 2 \ldots, n\}$, where each edge appears independently with probability p. We introduce a random binomial initial state, which we refer to as $\mathcal{S}_{1/2} \in \{0,1\}^{[n]}$. For every vertex $v \in V_n$, we have that $\mathbb{P}[\mathcal{S}_{1/2}(v) = 1] = \mathbb{P}[\mathcal{S}_{1/2}(v) = 0] = 1/2$, independently of any other vertex. We say that a sequence of events E_n defined on a sequence of probability spaces with probability measure $\mathbb{P}_n$ occurs *asymptotically almost surely* (or *a.a.s.*) if $\mathbb{P}_n[E_n] \to 1$ as $n \to \infty$. We now state our first result which describes the vertex strategies of an interacting node system on $G(n,p)$, with initial state $\mathcal{S}_{1/2}$. We denote this system by $(G(n,p), Q, \mathcal{S}_{1/2})$.

Theorem 1.1.1. *Let Q be a 2×2 non-degenerate payoff matrix. For any $\varepsilon \in (0,1]$ there exist positive constants Λ, n_0 such that for all $n \geq n_0$, if $p > \Lambda n^{-\frac{1}{2}}$, then with probability at least $1-\varepsilon$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/2}$, the interacting node system $(G(n,p), Q, \mathcal{S}_{1/2})$ will be unanimous after at most four rounds.*

While the above result only allows us to consider densities of $p \gg \Lambda n^{-1/2}$ we remark that placing a condition on Q allows us to consider densities just below the connectivity threshold of $G(n,p)$. For a 2×2 payoff Q with rows and columns indexed with $\{0,1\}$ we define the *payoff skew* of Q as,

$$\lambda = \lambda(Q) := (q_{1,1} - q_{0,1})/(q_{0,0} - q_{1,0}).$$

We will discuss the interpretation of λ (and its relationship to a global strategic bias) during the derivation of evolution rules, this is found within Section 2.1.1. For the following results, we assume that $\lambda \neq 1$. Under this assumption we are able to consider sparser values of p, which include passing well below the connectivity threshold to just above the critical threshold of $G(n,p)$. In the sub-critical regime, we can no longer hope for unanimity, as all components are isolated vertices or trees of size $O(\log n)$. The following result states for $p \gg 1/n$ almost all vertices will be unanimous with $O(\log n)$ rounds.

Theorem 1.1.2. *Let $p = d/n \leq 1$, where $d \gg 1$, and let Q be a 2×2 non-degenerate payoff matrix. Suppose that $(G(n,p), Q, \mathcal{S}_{1/2})$ is an interacting node system with payoff skew $\lambda \neq 1$. For any $\varepsilon > 0$ there exists $\beta = \beta(\lambda, \varepsilon) > 0$ such that a.a.s. at least $n(1-\varepsilon)$ vertices in $G(n, d/n)$ will be unanimous after at most $\beta \log n$ rounds.*

Moreover, there exists a constant $\alpha(\lambda) > 1$ such that if $d > \alpha(\lambda) \log n$, then a.a.s. $G(n, d/n)$ will be unanimous after one round.

The part of the above theorem for $np > c \log n$ with $c > \alpha(\lambda)$ implies that in fact, it is much stronger than Theorem 1.1.1. That is, for $\lambda \neq 1$ the system goes into unanimity in one round for much lower densities than $n^{-1/2}$. Thus, it will remain to prove Theorem 1.1.1 only for $\lambda = 1$.

We refine the above theorem focusing on the largest connected component of $G(n,p)$, which we denote as $L_1(G(n,p))$ (formally, if there are at least two, we take the lexicographically smallest one). Let $u_n^{(1)}$ be the probability that $L_1(G(n, d/n))$ will eventually become unanimous. The next two theorems give the precise location of the threshold on p above which $u_n^{(1)}$ approaches 1. In fact, there are two different thresholds for the two regimes. We start with the majority regime. For $\lambda > 0$, let

$$\ell_\lambda := \lceil \max\{\lambda, \lambda^{-1}\} \rceil \text{ and } c_\lambda := \frac{1}{\ell_\lambda + 1}.$$

(For a real number $x > 0$, we let $\lceil x \rceil = x$, if $x \in \mathbb{N}$, and $\lceil x \rceil = \lfloor x \rfloor + 1$, otherwise.) We write $\mathrm{Po}(\gamma)$ for the Poisson distribution with parameter $\gamma > 0$.

Theorem 1.1.3. *Suppose that $\lambda \neq 1$ and $d = c_\lambda \log n + \log\log n + \omega(n)$. Then the following hold in the majority regime.*

1. *If $\omega(n) \to +\infty$ as $n \to +\infty$, then*

$$\lim_{n\to+\infty} u_n^{(1)} = 1.$$

2. *If $\omega(n) \to c \in \mathbb{R}$ as $n \to +\infty$, then*

$$\limsup_{n\to+\infty} u_n^{(1)} \leq \sum_{k=0}^{\infty} \left(1 - \frac{1}{2^{\ell_\lambda+1}}\right)^k \mathbb{P}\left(\mathrm{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!) = k\right).$$

and

$$\liminf_{n\to+\infty} u_n^{(1)} \geq \sum_{k=0}^{\infty} \left(\frac{1}{2^{\ell_\lambda+1}}\right)^k \mathbb{P}\left(\mathrm{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!) = k\right).$$

3. *If $\omega(n) \to -\infty$ as $n \to +\infty$, then*

$$\lim_{n\to+\infty} u_n^{(1)} = 0.$$

The analogous result for the minority regime is as follows. For $\lambda > 0$, we let

$$\ell'_\lambda = \lfloor \max\{\lambda, \lambda^{-1}\} \rfloor.$$

Theorem 1.1.4. *Suppose that $\lambda \neq 1$ and $d = \frac{1}{2}\log n + \frac{1+\ell'_\lambda}{2}\log\log n + \omega(n)$. Then the following hold in the minority regime.*

1. *If $\omega(n) \to +\infty$ as $n \to +\infty$, then*

$$\lim_{n\to+\infty} u_n^{(1)} = 1.$$

2. If $\omega(n) \to c \in \mathbb{R}$ as $n \to +\infty$, then

$$\limsup_{n\to+\infty} u_n^{(1)} \leq \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k \mathbb{P}\left(\mathrm{Po}(e^{2c}/\ell'_\lambda!) = k\right).$$

and

$$\liminf_{n\to+\infty} u_n^{(1)} \geq \mathbb{P}\left(\mathrm{Po}(e^{2c}/\ell'_\lambda!) = 0\right).$$

3. If $\omega(n) \to -\infty$ as $n \to +\infty$, then

$$\lim_{n\to+\infty} u_n^{(1)} = 0.$$

The reason for the existence of two different thresholds is that the structures that block unanimity are different in the two regimes. Effectively, these are the thresholds for their disappearance as subgraphs of $G(n,p)$ and, as it turns out, as subgraphs of $L_1(G(n,p))$. Nevertheless, as our arguments show, in both regimes the unanimity of $L_1(G(n,p))$ is achieved in $O(\log n)$ steps.

During our analysis, we will see that there are two different kinds of unanimity depending on whether the payoff matrix is in the majority or in the minority regime. In the majority regime, all vertices of $L_1(G(n,p))$ stabilise to one of the two strategies. However, in the minority regime, the vertices of $L_1(G(n,p))$ arrive at unanimity but they fluctuate incessantly between the two strategies. In other words, in the majority regime the subsystem of $L_1(G(n,p))$ becomes periodic with period 1, whereas, in the minority regime, the period is equal to 2. Furthermore, we identify the strategy that is played at each step once this subsystem has entered the periodic cycle.

1.2 Best Response Dynamics with More than Two Strategies

We consider an extension to the work given in Section 1.1. In this section, we are concerned with the evolution of best response dynamics on $G(n, p)$ where each agent may possess any constant $\ell \in \mathbb{N}$ number of strategies. We deduce a classification for the evolution of agent behaviour. This classification depends on a collection of terms derived from the payoff matrix, analogous to the λ term in the previous section. We consider two possible cases that may arise depending on the value of this term and deduce unanimity in each of these settings. We also suggest future work and approaches, in order to analyse the behaviour of the system in the remaining cases.

1.2.1 Best Response Dynamics with ℓ Strategies

Consider an $\ell \times \ell$ payoff matrix $Q = [q_{i,j}]_{0 \leq i,j \leq \ell-1}$, with each $q_{i,j} \in \mathbb{R}$. Each player will now simultaneously choose a strategy from $\{0, 1, \ldots, \ell - 1\}$. We recall that player 1 will then receive a payoff given by $q_{i,j}$, and player 2 will receive a payoff $q_{j,i}$. Again, we work on a fixed graph $G = (V, E)$ where if two nodes are adjacent, then the corresponding agents interact with each other by playing the matrix game Q. Again the system has the same setup as in Section 1.1.3. We fix a graph G, a payoff matrix Q, and for every $v \in V(G)$ an initial vertex strategy $S_0(v)$ where $S_0 : V(G) \to \{0, 1, \ldots, \ell - 1\}$. Again the process evolves in discrete time. We recall for each $t \in \mathbb{N}_0$ that $S_t(v)$ denotes the strategy played by vertex v at step t. We have that $n_t(v; j) = |N_G(v) \cap \{u : S_t(u) = j\}|$. Each vertex will observe the strategies of all of its neighbours and compute the total payoff it would have received had it played each of the strategies. We can summarise this information by the *score vector* of v at time t, which we define as $T_t(v) = (T_t(v; 0), T_t(v; 1), \ldots, T_t(v; \ell - 1))$

with

$$T_t(v) := \begin{bmatrix} q_{0,0} & \cdots & q_{0,\ell-1} \\ \vdots & \ddots & \vdots \\ q_{\ell-1,0} & \cdots & q_{\ell-1,\ell-1} \end{bmatrix} \cdot \begin{bmatrix} n_t(v;0) \\ \vdots \\ n_t(v;\ell-1) \end{bmatrix}.$$

To progress from round t to round $t+1$ we consider the row(s) containing the maximum score in the vector $T_t(v)$. In the case that

$$\underset{0\le i\le \ell-1}{\operatorname{argmax}}\left(T_t(v;i)\right) = \{i \in \{0,\dots,\ell-1\} : T_t(v;i) \ge T_t(v;j) \text{ for all } j \in \{0,1,\dots,\ell-1\}\}$$

is uniquely defined, we have that

$$S_{t+1}(v) = \underset{0\le i\le \ell-1}{\operatorname{argmax}}\left(T_t(v;i)\right).$$

In the case that the above term is not unique, equivalently there exist at least two rows which achieve the maximum value in $T_t(v)$, we have two cases. If we have that $S_t(v) \in \operatorname{argmax}_{0\le i\le \ell-1}(T_t(v;i))$ then $S_{t+1}(v) = S_t(v)$. Otherwise, it follows that $S_{t+1}(v)$ is a uniform random variable over $\operatorname{argmax}_{0\le i\le \ell-1}(T_t(v;i))$. In practice, these tiebreak situations have little impact on the subsequent analysis. We justify such rules in an informal sense. If an agent's current strategy is among the best total payoffs, then the agent has no incentive to switch to another strategy. Otherwise, if there exists a set of equally optimal strategies different from $S_t(v)$, then the agent has the incentive to switch. As these strategies are equally rewarding, the agent should have no preference for any single strategy. As a result, the agent chooses from this set uniformly at random.

1.2.2 Main Results for Best Response Dynamics with $\ell > 2$ Strategies.

In this section, we state our two main theorems. Both theorems declare the unanimity of the system for a given density and a condition on the matrix Q. We recall that the system possesses an underlying topology given by the binomial random graph, $G(n,p)$.

We denote the random multinomial initial state on $\ell \in \mathbb{N}$ strategies as, $\mathcal{S}_{1/\ell}$, wherein for all $k \in \{0, 1, \ldots \ell - 1\}$ we have

$$\mathbb{P}[S_0(v) = k] = 1/\ell$$

for all $v \in V(G)$. Furthermore, our analysis depends heavily on the sums of each row in Q. We introduce some notation to denote the number of rows in Q that achieve the maximum row sum. For an $\ell \times \ell$ payoff matrix Q and each $i \in \{0, 1, \ldots, \ell - 1\}$ we denote,

$$\Sigma R_i = \sum_{k=0}^{\ell-1} q_{i,k},$$

to be the sum of the entries in row i. We will be interested in how many rows achieve a maximum row sum in Q. Thus we define,

$$M(Q) := \left|\operatorname*{argmax}_{0 \leq i \leq \ell-1}\{\Sigma R_i\}\right| = |\{i \in \{0, 1, \ldots, \ell-1\} : \Sigma R_i \geq \Sigma R_j \text{ for all } j \in \{0, 1, \ldots, \ell-1\}\}|.$$

Our two theorems consider the case where $|M(Q)| \in \{1, 2\}$. Furthermore, throughout the remainder of our analysis we demand that Q has unique column maxima, i.e., for every $j \in \{0, 1, \ldots \ell - 1\}$ we have

$$\left|\operatorname*{argmax}_{0 \leq i \leq \ell-1}\{q_{i,j}\}\right| = |\{i \in \{0, 1, \ell - 1\} : q_{i,j} \geq q_{k,j} \text{ for all } \{0, 1, \ldots, \ell - 1\}\}| = 1.$$

This condition is required for the system to stay unanimous after unanimity is reached for the first time. This follows from Lemma 1.2.3, as the evolution of a unanimous system is dictated by the column maxima. If there exists a column with non-unique column maxima, then a totally unanimous system can potentially evolve by vertices making independent random choices. At this point, determining the long-term behaviour of the system becomes intractable.

Our two theorems are distinguished by the value of $M(Q)$. In both cases, the system rapidly achieves unanimity. In the case that $M(Q) = 1$, or equivalently there is a unique maximum row sum, unanimity is achieved in 1 round a.a.s.

Theorem 1.2.1. *Let p be such that, $p \gg n^{-\alpha}$, for $\alpha \in (0,1)$. Suppose $\ell \in \mathbb{N}$ and Q is an $\ell \times \ell$ payoff matrix. Furthermore suppose $M(Q) = 1$, $i^* = \mathrm{argmax}_{0 \le i \le \ell-1}\{\Sigma R_i\}$, and the column maxima of Q are uniquely defined. Then across the product space of $\mathcal{S}_{1/\ell}$ and $G = G(n,p)$ we have with probability $1 - o(1)$ that $S_1(v) = i^*$ for every $v \in V_n$.*

We consider the case where $M(Q) = 2$, equivalently there exists exactly two rows with equal maximal row sums. We consider the case where $\ell = 3$, however our arguments can be adapted with minimal changes to prove an analogous result for an $\ell \times \ell$ matrix. In the initial state, the strategy with the minimal row sum covers $o(1)$ vertices in the following round. For the remaining vertices, the behaviour of the first round is governed by a quantity which depends on the entries of Q and the global distribution of strategies. This quantity determines which of the two remaining strategies receives a boost in the following round. The unique column maxima requirement ensures a bias exists between these two strategies. In the following round, the entries of Q determine the evolution of a large number of vertices. The unique column maxima condition ensures one of the three strategies is identified as a leader, and this results in a large bias towards this strategy. Finally, the column maxima alone will then determine which strategy is first entered as the system reaches unanimity.

Theorem 1.2.2. *Let $np \gg n^{2/3}$. Suppose that Q is a 3×3 payoff matrix, with $M(Q) = 2$ and a unique column maxima. Then for every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that for all $n > n_0$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/3}$, we have that there exists some $i^* \in \{0,1,2\}$ such that*

$$\mathbb{P}\left[\bigcap_{v \in V_n} \{S_3(v) = i^*\}\right] \ge 1 - \varepsilon.$$

Each of the above theorems declares an upper bound for when each system first reaches unanimity. The following lemma dictates the behaviour of the system beyond this point for any payoff matrix Q with unique column maxima. If all vertices are playing strategy

k, they will evolve by playing the strategy contained in the row of column k's maximum. Given that the unique column maxima condition is met, this process is always deterministic.

Lemma 1.2.3. *Suppose G contains no isolated vertices, and Q an $\ell \times \ell$ payoff matrix with distinct column maxima. Suppose that for all $v \in V(G)$ we have that $S_t(v) = k$ for some strategy k. Then we have that $S_{t+1}(v) = \text{argmax}_{0 \le i < \ell - 1}\{q_{i,k}\}$.*

Proof. The proof of this lemma follows by direct calculation of the score vector for each vertex v. As all vertices have at least one neighbour and the system is unanimous, we note there exists some k such that every vertex has $n_t(v; j) = 0$ for $j \neq k$ and $n_t(v; k) > 0$. Hence for the vector $\mathbf{n}_t(v) = (n_t(v; 0), n_t(v; 1), \ldots, n_t(v; \ell - 1))$ we have that,

$$T_t(v) = Q \cdot \mathbf{n}_t(v) = n_t(v; k)\mathbf{C}_k,$$

where $\mathbf{C}_k$ is the k^{th} column of Q. As Q has unique column maxima, we have that $\mathbf{C}_k$ has a unique maximum entry in some row k'. Thus we have that $S_{t+1}(v) = k' = \text{argmax}_{0 \le i < \ell}\{q_{i,k}\}$.

□

In light of Lemma 1.2.3, it is of interest to consider what long-term behaviour can occur. Firstly, if vertices are playing a strategy k and the column maximum of column k is also in row k then the system stabilises. Thus column maxima on the leading diagonal can lead to stabilisation in that strategy. Consequently, matrices which have all column maxima off the leading diagonal can never stabilise under the above hypothesis. If the system does not stabilise then the vertices will unanimously cycle through a subset of the strategies.

Finally, it is of note to consider the case when Q does not possess unique column maxima. Suppose the vertices are unanimously playing strategy k, and $\mathbf{C}_k$ contains multiple column maxima, lying off the leading diagonal. It follows that each vertex will choose one of these strategies independently and uniformly at random. This leads to a situation where unanimity is lost entirely and we essentially have a new initial global distribution from a subset of the strategies. At this point, deducing any long-term behaviour is largely infeasible, if any discernible long-term behaviour exists at all.

1.3 The Modularity of Random Graphs on the Hyperbolic Plane

The final result of this thesis concerns the computation of a prominent network metric on a certain model of complex networks. In this section, we provide a brief overview of the rich and emergent field of complex networks, along with a brief description of the popular metrics used for analysis. We conclude this survey with a more in-depth overview of our metric of interest, *modularity*. We then proceed to define the Hyperbolic Random Graph model and state our central result concerning the modularity score of these networks, namely Theorem 1.3.1

1.3.1 Complex Networks

A wide array of complex physical phenomena can be identified by studying how smaller parts of the system are connected. For example, the complex information transfer within the internet can be studied through the lens of many physically connected computers. Similarly, analysing how diseases propagate globally through populations can be understood by having a grasp on both the structure of local communities and how these communities connect.

Complex networks are generally considered to be networks which arise as representations for some sufficiently large real-world systems or phenomena. This notion can feel generally quite imprecise, as different systems can encode wildly different networks; however, we can broadly classify these networks by the idea that they possess similarly valued network theoretic metrics.

1.3.2 Properties of Complex Networks

There are a variety of properties and metrics which are desirable in classifying complex networks. We outline a number of common metrics: *Small Worlds*; *Scale-Free Degree Distribution*; *Community Structure*; and *Modularity*. We direct the reader to [6, 26, 71] for consideration of other prominent metrics; including, *Clustering Coefficients*, *Motifs*, and *Centrality*.

Small World Phenomena

Despite the large size of a complex network, any pair of vertices are typically linked by short paths. This is a broadly similar notion to observing that these networks possess small diameter. This apparent property was coined by Milgram [61] as the *small world* effect. In the context of the population networks, this property is sometimes popularly referred to as the "six degrees of separation", wherein any pair of people can be linked together by a chain of at most six acquaintances. In more recreational settings this inspired the suggestion of the Erdős number, defined to be the distance a given mathematician is from Paul Erdős in the co-authorship collaboration graph, see [63]. In popular culture, even more variants of this quantity exist. The most well-known of these is the Bacon number, the distance an actor is from Kevin Bacon based on mutual film credits. Both of these quantities form crude upper bounds on both the radius and diameter of these networks. Among those with a finite Bacon number, the average is 2.9994, indicating positive empirical evidence for the small world properties of this network [10].

Mathematical models of complex networks generally display a diameter of polylogarithmic order as a function of the number of vertices, see [36]. This is explained by considering the degree distribution: a small number of nodes appear to have an atypically large degree. As a consequence, many of the shortest paths end up passing through a small collection of "highly popular" nodes. This was empirically suggested during Milgram's letter experiment [61]. Milgram tasked a diverse selection of participants to send a letter to one target person in New York. The participants were only allowed to forward the letter to an immediate friend or acquaintance who was more likely to know the target person. Equivalently, Milgram had described an empirical variation of searching for a target vertex in a network. Milgram made two key observations of the completed chains: many were short (< 10) connections, and almost all chains passed through the same two acquaintances at the penultimate step. This later property appears to mimic the idea that in some complex networks, global connectivity is generally dependent on a small set of popular nodes. For example, the connectivity in some regimes of the hyperbolic random graph model can be determined by the presence or absence of nodes near the centre [14].

Scale Free Networks and Power Law Degree Distributions

As remarked above, the small-world phenomena is primarily driven by a small number of nodes with a large degree. Conversely, many of the remaining nodes in the network have a significantly smaller degree. We quantify this as follows: for a given network and any integer $k \geq 0$, denote p_k to be the proportion of nodes with degree k. We say a network is *scale free* if there exists positive constants C and γ such that $p_k = Ck^{-\gamma}$, we remark that such a p_k is said to follow a *power-law* distribution.

The scale of the degree distribution is dependent on the method of generation. Some of the earliest constructions of networks with a scale-free degree distribution were formulated by Barabási and Albert [3]. In their preferential attachment model, they considered an evolving process, where new vertices join the network and attach to a subset of existing vertices with probabilities proportional to the vertex's current degree. This leads to a rich-get-richer style of generation, where a small number of nodes achieve a relatively large degree.

The below figure shows the comparison of the degree counts between two graph models: the former, a Barabási-Albert model where a newly arriving vertex connects to a single existing vertex with a distribution as above, while the latter depicts the standard Erdős-Rényi binomial random graph.

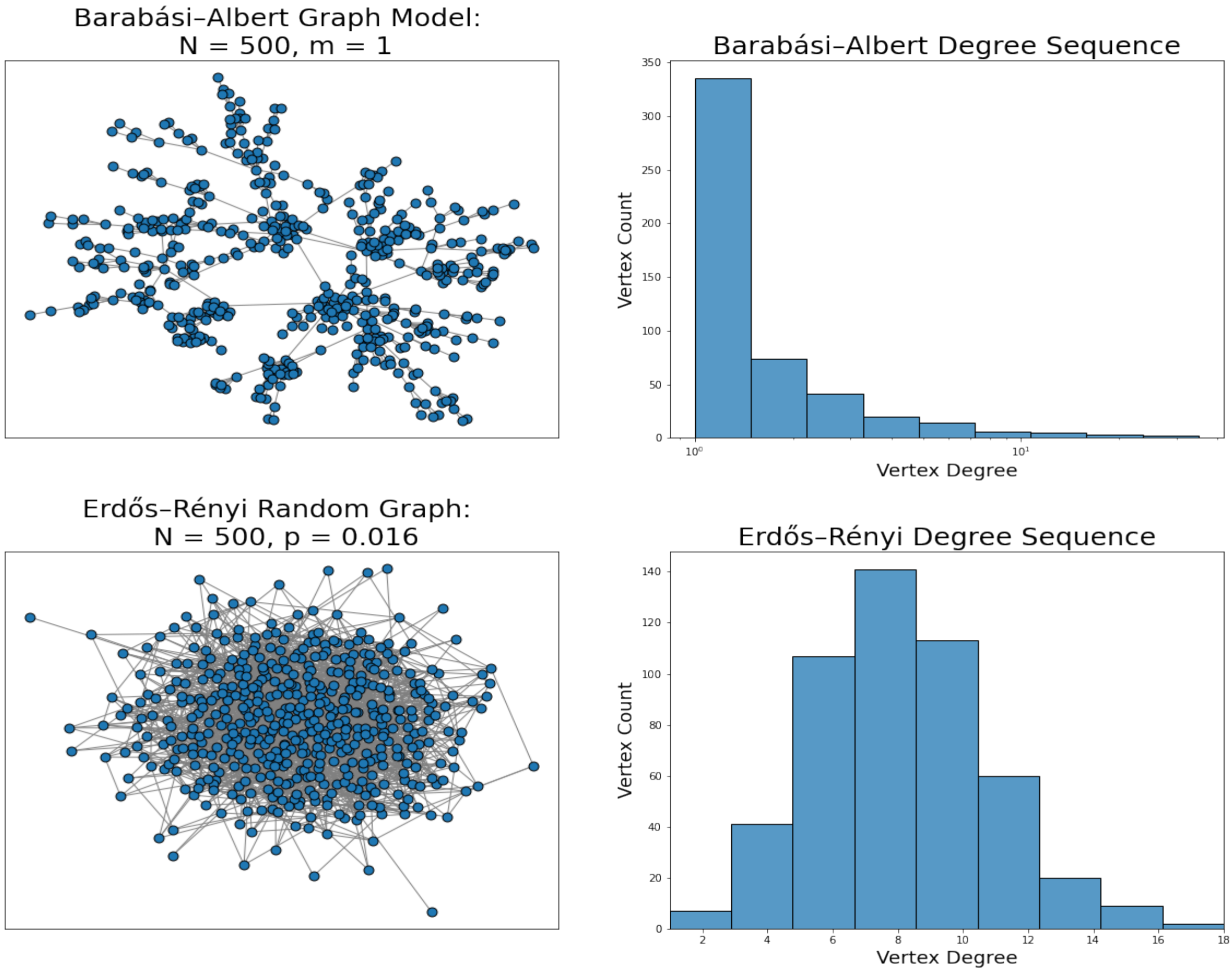

Figure 1.3: Comparison of the degree distribution between the Barabási-Albert and Erdős-Rényi Graph Models. Diagrams are produced with the Python NetworkX package, [42].

The Barabási-Albert preferential attachment model displays a clear exponential decay for the number of vertices showing a larger degree. In contrast, the vertices in the Erdős-Rényi random graph possesses degrees which are independently and identically distributed binomial random variables. Consequently, the degree distribution roughly obeys a Gaussian limit law centred on np. Therefore, in contrast to the scale-free preferential attachment network, we say that the Erdős-Rényi random graph possesses a single-scale, see [31].

Community Structure and Modularity

M. Granovetter, in his pioneering analysis of social networks [40], pointed out that a fundamental feature of social networks is the distinction between weak links and strong links. These strengths reflect the intensity of interaction between two individuals, which may be dependent on measures such as the frequency of interaction. As Granovetter remarked, an individual is more likely to interact with other individuals through strong links. This is expressed in terms of structural features of the social network, whereby individuals belong to communities, tightly knit by strong links, and these communities are typically joined through weak links. These ideas postulate that a fundamental characteristic of social networks is the existence of communities, or modules, within such a network. These are mutually disjoint subsets of nodes/individuals which have many edges connecting them but are joined to other modules by few edges.

The applications of community detection can be wide-ranging. Primarily, if we view the communities as induced subgraphs, then they possess broadly different properties compared to the network globally. On average, communities display higher local clustering and edge density compared to the entire network. Consequently, this can be exploited to detect the communities themselves, see [44]. Other applications can be seen in [24], whereby assuming an underlying community structure can be used to detect and correct errors within the data collection stages of encoding real-world networks.

Identifying such a partition in a given social network, or any other complex network is computationally challenging [16]. But before we set out to find algorithms that give even an approximate solution to this problem, one needs to quantify what defines a quality partition capturing the underlying community structure of the network. Such a quantification was given by Newman and Girvan [72] and is called the *modularity score* of a given partition.

For a graph G with $m \geq 1$ edges, define the *modularity score* associated with the partition $\mathcal{A}$ of the vertex set V to be

$$\mathrm{mod}_{\mathcal{A}}(G) = \sum_{A \in \mathcal{A}} \left(\frac{e(A)}{m} - \left(\frac{\mathrm{vol}(A)}{2m} \right)^2 \right)$$

where $e(A)$ denotes the number of edges within part A and $\mathrm{vol}(A) = \sum_{v \in A} \deg(v)$ denotes the volume of A, that is, the sum of the degrees of the vertices in A.

For graphs G without edges define $\mathrm{mod}_{\mathcal{A}}(G) = 0$. Note that the definition of modularity extends naturally to weighted graphs and is often used in the weighted form in applications. The term $e(A)$ becomes the sum of the weights of edges in A and the degree of a vertex $\deg(v)$ is the sum of the weights of the edges incident to v. The summation that defines $\mathrm{mod}_{\mathcal{A}}(G)$ is a comparison between the given network G and a random network with the same degree sequence. The first term $\frac{1}{m} \sum_{A \in \mathcal{A}} e(A)$ is the probability that a randomly chosen edge of G will lie inside one of the parts, whereas the term $\sum_{A \in \mathcal{A}} (\mathrm{vol}(A)/2m)^2$ represents the probability that a random edge lies in one of the parts in a uniformly random graph with the same degree distribution as G. On one extreme if there were no edges between the parts of $\mathcal{A}$, then $\frac{1}{m} \sum_{A \in \mathcal{A}} e(A) = 1$. If $\mathcal{A}$ consists of a large number of parts that are comparable in volume, then the second term $\sum_{A \in \mathcal{A}} (\mathrm{vol}(A)/2m)^2$ is small. Hence, such a highly modular partition will have a modularity score close to 1. With $\mathcal{P}(V)$ denoting the set of all partitions of V the *modularity of graph G* is then

$$\mathrm{mod}(G) = \max\{\mathrm{mod}_{\mathcal{A}}(G) \ : \ \mathcal{A} \in \mathcal{P}(V)\}.$$

The set $\mathcal{P}(V)$ includes the trivial partition $\{V\}$ placing all vertices into the same part. Note that the modularity score of $\{V\}$ is zero for any graph. Hence for any graph $0 \leq \mathrm{mod}(G) < 1$ with values near 1 taken to indicate a high level of community structure and values near 0 taken to indicate a lack of community structure.

The diagram below shows the difference in modularity score achieved by two vertex partitions. The former shows a vertex partition focused on maximising modularity. As a result, this partition contains densely clustered communities, with weak connections between them. The latter partition is chosen randomly, but with the same number of parts as the previous algorithm. Clearly, this partition fails to capture any of the underlying community structure and consequently achieves a modularity score close to zero.

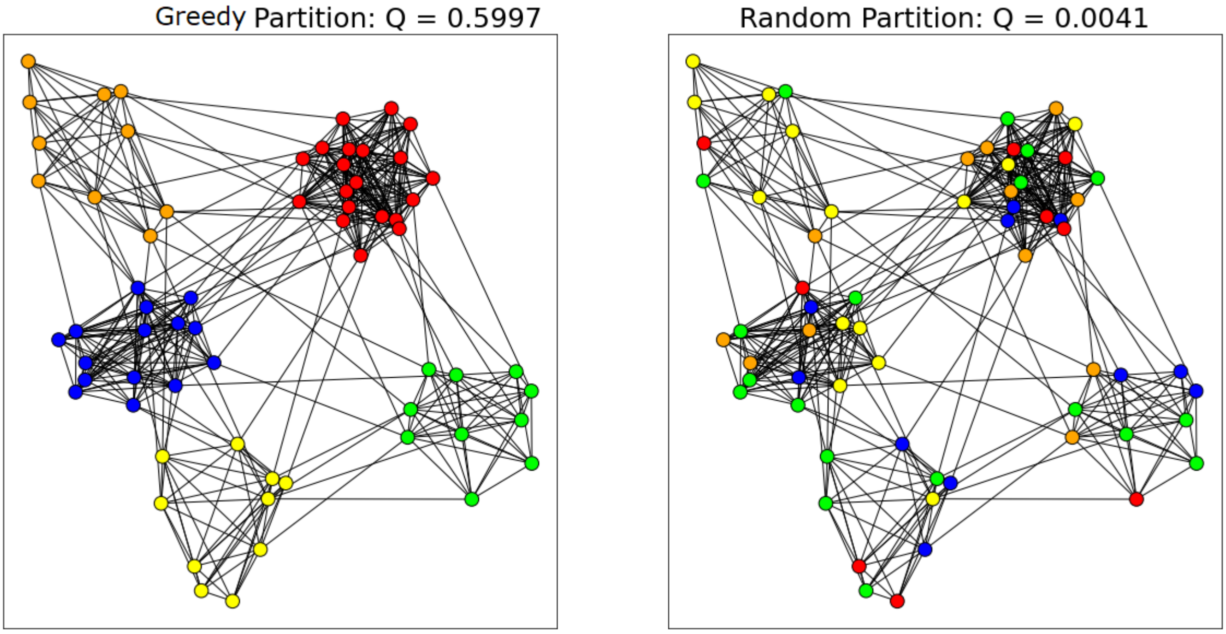

Figure 1.4: A comparison of two vertex partitions, with parts indicated by vertex colour. The former partition is found using the greedy Clauset-Newman-Moore algorithm, see [25]. The latter is a random vertex partition with the same number of parts. The associated modularity score is given by Q. Diagrams produced with the Python NetworkX package, [42].

1.3.3 The Hyperbolic Random Graph

Krioukov et al. [50] introduced a model of random geometric graphs on the hyperbolic plane as a model of complex networks, which we abbreviate as the *KPKBV model* after its inventors. This is based on the assumption that the geometry of the hyperbolic plane can accommodate the hidden hierarchy of a complex network and its intrinsic inhomogeneity. Their basic assumption is that the hierarchies that are present in a complex network induce a tree-like structure, and this suggests that there is an underlying geometry of a complex network which is hyperbolic.

There are several representations of the standard hyperbolic plane H^2_{-1} of curvature -1. We shall use the Poincaré unit disc representation, which is simply the open disc of radius one, that is, $\{(u,v) \in \mathbb{R}^2 \ : \ u^2+v^2<1\}$, which is equipped with the hyperbolic metric: $4\,\frac{du^2+dv^2}{(1-u^2-v^2)^2}$. This is a standard formulation of the hyperbolic plane. In particular, a suitable integration of the metric shows that the length of a circle of (hyperbolic) radius r (centred at the origin) is $2\pi\ \sinh(r)$, whereas the area of this circle (centred at the origin) is $2\pi(\cosh(r)-1)$. Hence, a fundamental difference with the Euclidean plane is that volumes grow exponentially.

The KPKBV model introduced by Krioukov et al. [50] yields a random geometric graph on H^2_{-1}. Consider the Poincaré disc representation of the hyperbolic plane H^2_{-1}. The random graph will have n vertices, we take all asymptotics with respect to n. Let $\nu > 0$ be a fixed constant and let $R = R(n) > 0$ satisfy $n = \nu e^{R/2}$.

Consider the disc $\mathcal{D}_R$ of hyperbolic radius R centred at the origin of the Poincaré disc (that is, the set of points of the Poincaré disc at a hyperbolic distance of at most R from its origin).

We take a random set of points of size n that are the outcomes of the *i.i.d.* random variables $v_1,\ldots,v_n$ taking values on $\mathcal{D}_R$. (We will be referring to the random variables v_i as vertices, meaning their values on $\mathcal{D}_R$.) More specifically, assume that v_1 has *polar* coordinates (r,θ). The angle θ is uniformly distributed in $(0,2\pi]$ and the probability density function of r, which we denote by $\rho_n(r)$, is determined by a parameter $\alpha > 0$ and is equal to

$$\rho_n(r) = \begin{cases} \alpha\frac{\sinh(\alpha r)}{\cosh(\alpha R)-1} & \text{if } 0 \le r \le R; \\ 0 & \text{otherwise.} \end{cases} \tag{1.2}$$

The aforementioned formulae for the area and the length of a circle of a given radius imply that if we set $\alpha = 1$, the distribution described in (1.2) is the uniform distribution on $\mathcal{D}_R$ (under the hyperbolic metric). For general $\alpha > 0$ Krioukov et al. [50] called this the *quasi-uniform* distribution on $\mathcal{D}_R$. Let us remark that in fact this is the uniform distribution on a disc of hyperbolic radius R within $H^2_{-\alpha^2}$ (the hyperbolic plane that has curvature $-\alpha^2$).

Given the point process $V_n = \{v_1, \ldots, v_n\}$ on $\mathcal{D}_R \subset H^2_{-1}$ and the fixed parameters α and ν we define the random graph $\mathcal{G}(n; \alpha, \nu)$ on the point-set of V_n, where two distinct points form an edge if and only if they are within (hyperbolic) distance R from each other. Figure 1.5 (below) shows the ball of radius R around a point $p \in \mathcal{D}_R$, denoted by $B(p; R)$. Thus, any point/vertex of $\mathcal{G}(n; \alpha, \nu)$ that falls inside the shaded region becomes connected to p.

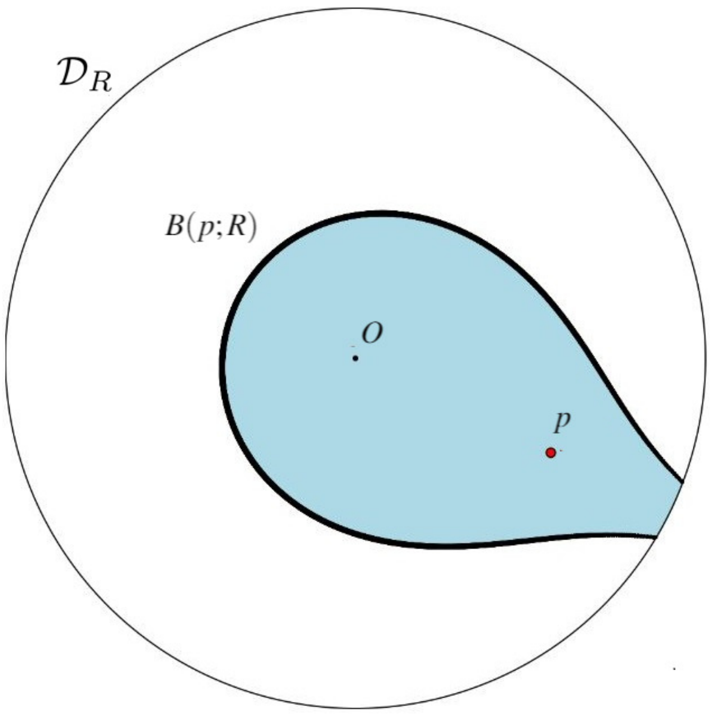

Figure 1.5: The ball of radius R centered at point p, within $\mathcal{D}_R$

We refer the discussion of the general network properties of the KPKBV model to Section 4.1.1. Below, we attach some realisations of this graph for various values of α.

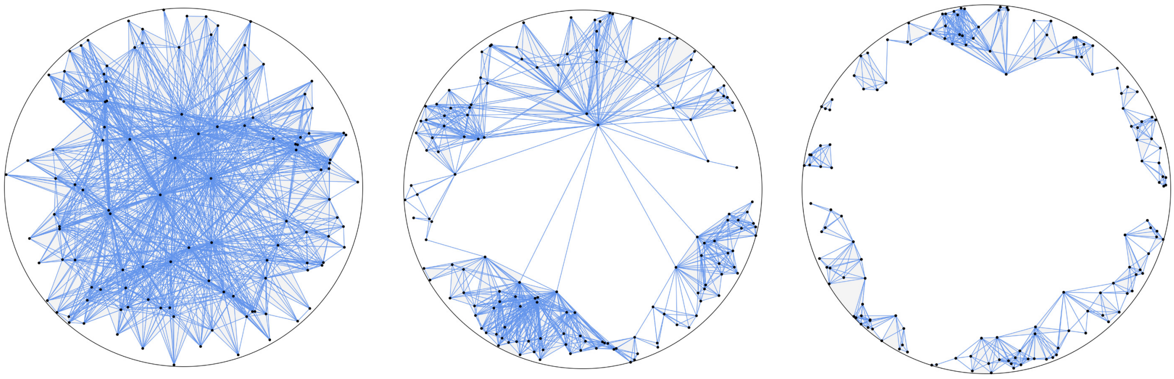

Figure 1.6: Three samples of $\mathcal{G}(n;\alpha,\nu)$ with a fixed $n = 150$, $\nu = 2$ and variable α. From left to right: $\alpha = 0.6$, $\alpha = 1$ and $\alpha = 1.8$

1.3.4 Poissonisation of the KPKBV Model and Main Results

We will work on the Poissonisation of the above model. Recall that $\mathcal{D}_R$ was defined to be the disc of hyperbolic radius R around the origin O of the Poincaré disc representation of the hyperbolic plane of curvature -1. Here, the vertex set is the point set of a Poisson point process on $\mathcal{D}_R$ with intensity

$$n\frac{1}{2\pi}\rho_n(r)drd\theta.$$

We denote it by $\mathrm{P}_{\alpha,\nu,n}$. We also denote by $\kappa_{\alpha,\nu,n}$ the Borel measure on $\mathcal{D}_R$ given by

$$\kappa_{\alpha,\nu,n}(S) = \frac{1}{2\pi}\int_S \rho_n(r)drd\theta,$$

for any Borel-measurable set S. Hence, the number of points that $\mathrm{P}_{\alpha,\nu,n}$ has inside S is distributed as $\mathrm{Po}\left(n \cdot \kappa_{\alpha,\nu,n}(S)\right)$. Moreover, the numbers of points in any finite collection of pairwise disjoint Borel-measurable subsets of $\mathcal{D}_R$ are independent Poisson-distributed random variables.

We will define the random graph whose vertex set is the set of points of $\mathrm{P}_{\alpha,\nu,n}$ in $\mathcal{D}_R$. As in $\mathcal{G}(n;\alpha,\nu)$, two vertices/points of $\mathrm{P}_{\alpha,\nu,n}$ are adjacent if and only if their hyperbolic distance is at most R. We denote the resulting graph by $\mathcal{P}(n;\alpha,\nu)$. The main theorem of this chapter is that with high probability the modularity of $\mathcal{P}(n;\alpha,\nu)$ is close to 1.

Theorem 1.3.1. *For any $\alpha > 1/2$ and $\nu > 0$, we have*

$$\text{mod}(\mathcal{P}(n;\alpha,\nu)) \to 1,$$

as $n \to \infty$, in probability.

The modularity of $\mathcal{P}(n;\alpha,\nu)$ approaches 1 as $n \to \infty$, without any dependence on the average degree or the existence of a giant component. Very recently, Kovács and Palla consider an empirical approach to community structure in hyperbolic networks, [49]. In particular, they show that almost optimal modularity scores (greater than 0.95) can be achieved, using simulations of the hyperbolic random graph; however, they work in a range of parameters where the resulting random graph exhibits a high clustering coefficient and large α. In light of this, our results show that for sufficiently large networks this dependency on parameters is no longer required to achieve optimal modularity.

1.4 General Terminology

For a graph G, the vertex set will be denoted by $V(G)$ or in the case that $V(G) = [n] = \{1, 2, \dots, n\}$, we may denote this by V_n. We denote $E(G)$ to be the edge set of G, while $e(G)$ represents the quantity $|E(G)|$, and denotes the number of edges in G. For two distinct vertices $v, u \in V$ we write $v \sim u$ to indicate that they are adjacent.

For a vertex v in a given graph, we denote $N_G(v) = \{u \in V(G) : u \sim v\}$, to be the neighbourhood of v in G. Furthermore, we denote $d(v) := |N_G(v)|$ to be the degree of v within the graph G, we may also denote this as $\deg(v)$ and the particular graph we refer to will be made specific by the context. If $S \subset V$, we write $e(S)$ as the number of edges in G that have both endpoints in S. For $v \in V$, we denote $d_S(v)$ to be the degree of v inside S, i.e $|\{u \in S : u \sim v\}|$.

We recall that if $\mathcal{E}_n$ is an event on the probability space $(\Omega_n, \mathbb{P}_n, \mathcal{F}_n)$, for each $n \in \mathbb{N}$, we say that the sequence $\mathcal{E}_n$ *occurs asymptotically almost surely (a.a.s.)* if $\mathbb{P}_n(\mathcal{E}_n) \to 1$ as $n \to \infty$. In our context, we will be using the term *a.a.s.* for the sequence of probability spaces of the binomial random graphs $G(n, p)$ and the hyperbolic random graphs $\mathcal{P}(n; \alpha, \nu)$.

Let a_n, b_n be two sequences of positive real numbers. We write $a_n = \Theta(b_n)$, to denote that there are real numbers $c, C > 0$ such that $cb_n \leq a_n \leq Cb_n$, for all natural numbers n. We also write $a_n \sim b_n$ to denote that $a_n/b_n \to 1$, as $n \to \infty$.

We also use the symbol $\sim$ in the context of random variables. In particular, for a random variable X, we will write $X \sim \mathrm{Bin}(s, q)$ to indicate that the distribution of X is binomial with parameters $s \in \mathbb{N}$ and $q \in [0, 1]$. While for $\lambda > 0$ we say $X \sim \mathrm{Po}(\lambda)$ to denote that X is Poisson with parameter λ. For real μ and positive σ, we write $N \sim \mathrm{N}(\mu, \sigma)$ to say that N is a normally distributed random variable, with mean μ and variance σ. For a probability space $(\Omega, \mathcal{F})$ and an event $A \in \mathcal{F}$, we denote 1_A to be the indicator function of A. Specifically $1_A = 1$ if A occurs and 0 otherwise.

1.4.1 General Tools

In this section, we detail some common results from probability theory which are utilised throughout this thesis. We provide suitable references in the statement of each result.

The Chernoff Bound

Throughout our arguments, we will use the standard Chernoff bounds for the concentration of binomially distributed random variables. The inequality we will use follows from Theorem 2.1 in [45]. If X is a random variable such that $X \sim \mathrm{Bin}(N, q)$, then for any $\delta \in (0, 1)$ we have

$$\mathbb{P}(|X - Nq| \geq \delta Nq) \leq 2e^{-\delta^2 Nq/3}. \tag{1.3}$$

The FKG Inequality

We also require a classical correlation inequality, the *FKG inequality.* We state it in the context of the $G(n,p)$. We say that a graph property $\mathcal{P}$ is *non-decreasing* if, for graphs G, H on V_n, whenever $G \in \mathcal{P}$ and $G \subseteq H$ then $H \in \mathcal{P}$. Similarly we say that $\mathcal{P}$ is *non-increasing*, if whenever $H \notin \mathcal{P}$ and $H \subseteq G$, then $G \notin \mathcal{P}$ as well. We state the FKG inequality as follows.

Theorem 1.4.1 The FKG inequality [5] (page 91)**.** *Let $\mathcal{P}_1$ be an non-decreasing graph property and $\mathcal{P}_2$ be a non-increasing graph property. Then for the binomial random graph, $G(n,p)$, we have the following:*

$$\mathbb{P}[G(n,p) \in \mathcal{P}_1 \cap G(n,p) \in \mathcal{P}_2] \leq \mathbb{P}[G(n,p) \in \mathcal{P}_1] \cdot \mathbb{P}[G(n,p) \in \mathcal{P}_2].$$

Convergence to a Poisson Variable

We will also require a result concerning Poisson convergence, stated as Theorem 1.22 in [15]. For $r \in \mathbb{N}$ we define the rth factorial moment of a random variable X, as $X^{(r)} = X(X-1)(X-2)\ldots(X-r+1)$.

Theorem 1.4.2 [15] (page 25)**.** *Let $\lambda = \lambda(n)$ be a non-negative, bounded and natural-valued function. We define the random variable $P_\lambda \sim \mathrm{Po}(\lambda)$, and let X_n be a sequence of integer-valued random variables. Suppose for all $r \in \mathbb{N}$ we have,*

$$\lim_{n\to\infty} \mathbb{E}\left[X_n^{(r)}\right] = \lambda^r.$$

Then we have that, $X_n \xrightarrow{d} P_\lambda$.

CHAPTER 2

BEST RESPONSE DYNAMICS ON RANDOM GRAPHS WITH BINARY STRATEGIES

In this chapter we consider the analysis of best response dynamics on $G(n,p)$, where the vertices can play either strategies 0 or 1. We lead with a survey of work in the field of evolutionary game theory, with a particular focus on games with an underlying topology. We will then briefly restate our main theorems, alongside a proof outline, before progressing into a full proof of each result.

2.1 Evolving Games with an Underlying Topology

The need for a dynamic game theory was pointed out by von Neumann and Morgenstern in their seminal book [85] which set the foundations of modern game theory. Research on dynamic games on populations was stimulated by settings in evolutionary biology. In 1973 John Maynard Smith and George Price [58] set the foundations of *evolutionary game theory*, trying to explain phenomena that arise in animal fighting, which contradict the traditional Darwinian theory.

Considering two types of behaviour: *hawk* and *dove*. Any two interacting members of the population play the well-known hawk-dove game (see Section 2.1.1 for a precise description of its payoff matrix). Maynard Smith and Price use this game to illustrate that in a population in which individuals may have one of the two types, it is better to have a mixture of hawks and doves, rather than a pure population consisting of only either hawks or doves. In this model, every member of the population may interact with any other member. In graph-theoretic terms, this is an evolutionary system on a complete graph.

Later on, Nowak and May [73] considered settings with non-trivial underlying topology. They considered a population of agents that are located on the vertices of a 2-dimensional lattice, in which every agent interacts only with its four neighbours. The interaction is that any two adjacent agents play the prisoner's dilemma (which we describe in section 2.1.1) in which they may *cooperate* or *defect*. The dynamics that were considered there depends on the total payoff of each agent that is accumulated by the four games it plays with its neighbours (or three or two, if it is located on the boundary of the lattice). The agents update their strategies synchronously, with each agent adopting the strategy of a neighbour who had the largest total payoff in the previous round. The main observation in [73] is the *co-existence* of the two strategies in the long term, despite the fact that the defect strategy is a *Nash equilibrium* for this game: that is, in a game between two players who both defect none has an interest in cooperating, given that the other does not. Similar dynamics were studied by Santos and Pacheko [83] in the preferential attachment model, which typically yields graphs which have some properties of complex networks.

A form of local best response dynamics was first considered by Gilboa and Matsui [38] in a continuous-time setting and for a population with no underlying topology, where everyone interacts with everyone else. Gilboa and Matsui showed the existence of *cyclically stable sets*. Roughly speaking, these are sets of configurations of the population in which the best response dynamics is trapped.

We recall that our main focus is the study of the evolution of best response dynamics on a random graph, through the model of an interacting node system. In our analysis of these systems, we show that these best response dynamics reduce to *generalised majority or minority dynamics*. These terms refer to a general class of discrete processes on graphs where vertices have two states and at each round, a vertex adopts the state of the majority or the minority of its neighbourhood, respectively. Majority games reward the individuals who follow strategies which are in line with the popular opinion; such games are clearly cooperative, and thus agents will tend to form large unanimous coalitions; see [43]. On the other hand, minority games capture the idea that agents will benefit from making choices that oppose the popular consensus. The idea of a minority game was introduced to characterise the behaviour of agents within the *El Farol Bar problem*. The problem describes a fixed population who will repeatedly attempt to synchronously choose their favourite evening location; however, only the individuals who choose less crowded locations will be rewarded [18,27]. This problem captures the underlying tension of minority games. Choosing the lucrative option will clearly reward agents with the greatest payoff; however, this payoff is easily spoiled if too many of the participants think alike. If agents adopt this line of thinking, then we deduce that the entire population will reject the optimal payoff, regardless of the fact that the most lucrative option will tend to be uncontested. In the case of the El Farol Bar problem, any deterministic strategy will ultimately fail to satisfy anyone.

In the context of market investments, majority and minority strategies are of central importance. As suggested in [57], traders utilising contrarian-like strategies will view the market functioning in a minority-game-like way. They believe goods on the market have a fundamental price and will invest in such a way as to discourage the market from deviating away from its fundamental value. On the other hand, investors who are *trend followers* view the market from the context of a majority game. These investors tend to inflate the price of goods which are already travelling on an upward trend. Such investment strategies are believed to be the central cause driving the *pricing bubble* phenomena [20,21,46]. The success of each of these strategies is heavily dependent on the payoff agents can expect to receive, along with the motives and actions of their rivals.

While the study of evolutionary games on populations with arbitrary mutual interactions is well established, a new focus has now been given to systems of agents which possess an underlying topology. Within these systems, agents are only able to interact with their topological neighbours. The topology of these interactions is commonly represented by means of an underlying network. Consideration of an underlying topology allows us to analyse how local decisions in the system can cascade out to form a global consensus. For example, in [55] it is shown that a small set of agents which oppose the current consensus, can cause a large contagion of opposing opinions to spread throughout the network. While in [12] systems of interacting agents on a lattice are considered, agents have the opportunity to asynchronously update their opinion at random times, which are driven by an underlying Poisson point process. In the windows of opportunity which are given by these random times, agents will choose to update their strategies to the one which will now give them the largest current payoff. It is shown that a steady state can be achieved within this system; the shape of this distribution is heavily dependent on how agents choose to update their opinion during these random times.

As previously mentioned, the underlying topology which is the main focus of this thesis is that of a binomial random graph $G(n, p)$. Our results show that in a wide range of densities, best response dynamics stabilises rapidly. We prove a general result which shows that this is achieved in at most 4 rounds when $np = \Omega(n^{1/2})$. However, if the game exhibits a certain form of bias in terms of its pure Nash equilibria, then we get very precise results on how these dynamics evolve for p such that $np \gg 1$. In the presence of such a bias, we determine a sharp threshold on p above which the largest connected component reaches consensus among its vertices. For p below this critical value, we identify those substructures inside the largest component that are in disagreement with the majority of the vertices therein. Hence, we are able to characterise the co-existence of strategies very precisely.

We lay out the remainder of this chapter as follows: In Section 2.1.1, we consider our first approaches to analysing the model, specifically we construct precise evolution rules which lead towards a classification of the payoff matrices. We follow with a consideration of the *degenerate payoff matrices* and show the behaviour of the system is generally resolved within a single round. Following this, We focus our attention on a specific class of payoff matrices where there exists a certain form of bias among the pure Nash equilibria of the game. This is analysed in Section 2.2. In this case, we identify a sharp threshold for p above which consensus is reached in the largest component of $G(n,p)$ in at most $\beta \log n$ rounds, for a positive constant β. This critical value is determined by the payoff matrix and is below the connectivity threshold of $G(n,p)$. Furthermore, we show that if $np > c \log n$, for some $c > 1$ which depends on the parameters of the system, then in fact the system reaches consensus only after one round. This is summarised by the restatement of the following theorem.

Theorem 1.1.2. *Let $p = d/n \leq 1$, where $d \gg 1$, and let Q be a 2×2 non-degenerate payoff matrix. Suppose that $(G(n,p), Q, \mathcal{S}_{1/2})$ is an interacting node system with payoff skew $\lambda \neq 1$. For any $\varepsilon > 0$ there exists $\beta = \beta(\lambda, \varepsilon) > 0$ such that a.a.s. at least $n(1-\varepsilon)$ vertices in $G(n, d/n)$ will be unanimous after at most $\beta \log n$ rounds.*

Moreover, there exists a constant $\alpha(\lambda) > 1$ such that if $d > \alpha(\lambda) \log n$, then a.a.s. $G(n, d/n)$ will be unanimous after one round.

In Section 2.3, we consider games where this bias is no longer present. In this scenario, we observe that the interacting node system reduces to the so-called *majority or minority dynamics* on $G(n,p)$. We sample a selection of results from [32], which concern the rapid stabilisation of agent strategies when playing a majority game on a dense random graph. Using these results as a basis, we proceed to prove our result concerning the formation of consensus strategies in the minority game analogue of the random graph majority game. By combining the above results, we may readily deduce that for any 2×2 non-degenerate real-valued payoff matrix and suitably dense random graph, the agents of the interacting node system will reach a consensus after at most four rounds, with high probability. We restate this result as follows.

Theorem 1.1.1. *Let Q be a 2×2 non-degenerate payoff matrix. For any $\varepsilon \in (0, 1]$ there exist positive constants Λ, n_0 such that for all $n \geq n_0$, if $p > \Lambda n^{-\frac{1}{2}}$, then with probability at least $1 - \varepsilon$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/2}$, the interacting node system $(G(n,p), Q, \mathcal{S}_{1/2})$ will be unanimous after at most four rounds.*

2.1.1 The Skew of the Payoff Matrix

Suppose a vertex v at time t has picked a strategy $S_t(v) = i$. From the above discussion, we observe that the *incentive* for v to switch strategy is the condition that $T_t(v) < T'_t(v)$. Expanding both terms and re-arranging, we observe the condition for a vertex playing strategy i to switch to strategy $1-i$, in the following round, is as follows:

$$n_t(v;0)\,(q_{i,0} - q_{1-i,0}) < n_t(v;1)\,(q_{1-i,1} - q_{i,1})\,. \tag{2.1}$$

We recall the definition of the payoff skew of the matrix Q as:

$$\lambda = \lambda(Q) := (q_{1,1} - q_{0,1})\,/\,(q_{0,0} - q_{1,0})\,.$$

We remark that $\lambda(Q)$ is positive and well-defined, if and only if, Q is non-degenerate. If Q is non-degenerate, there are two possible cases: Either $q_{0,0} > q_{1,0}$ and $q_{1,1} > q_{0,1}$; or we have that $q_{0,0} < q_{1,0}$ and $q_{1,1} < q_{0,1}$.

We lead with the former case. By substituting values of $i \in \{0,1\}$ into (2.1), we can rephrase the evolution conditions for each agent in terms of $n_t(v;0)$ and $n_t(v;1)$. Suppose $S_t(v) = 0$, then the incentive for changing strategy is given as follows:

$$n_t(v;0)\,(q_{0,0} - q_{1,0}) < n_t(v;1)\,(q_{1,1} - q_{0,1})\,.$$

We now re-arrange the above, and apply that $\lambda = (q_{1,1} - q_{0,1})/(q_{0,0} - q_{1,0})$, to form the first evolution rule for changing from $S_t(v) = 0$ to $S_{t+1}(v) = 1$:

$$n_t(v;0) < n_t(v;1)\frac{q_{1,1} - q_{0,1}}{q_{0,0} - q_{1,0}} = n_t(v;1)\lambda. \tag{2.2}$$

Similarly, if we instead have that $S_t(v) = 1$, then the incentive to change to zero in the next round is expressed as:

$$n_t(v;0)\left(q_{1,0}-q_{0,0}\right) < n_t(v;1)\left(q_{0,1}-q_{1,1}\right).$$

We recall that in this case we have $q_{0,1} < q_{1,1}$, and therefore we re-arrange as follows:

$$n_t(v;1) < n_t(v;0)\frac{q_{1,0}-q_{0,0}}{q_{0,1}-q_{1,1}} = n_t(v;0)\frac{q_{0,0}-q_{1,0}}{q_{1,1}-q_{0,1}} = n_t(v;0)\frac{1}{\lambda}. \tag{2.3}$$

Combing equations (2.2) and (2.3) provide the required evolution rules as described by equation (2.4). For the second case, where $q_{0,0} < q_{1,0}$ and $q_{0,1} > q_{1,1}$, an identical argument will produce the rules given by equation (2.5). If $S_t(v) = i$, and Q is such that $q_{0,0} > q_{1,0}$ and $q_{1,1} > q_{0,1}$, then we can write:

$$S_{t+1}(v) = \begin{cases} 1-i & \text{if } n_t(v;i) < \lambda^{1-2i} n_t\left(v;1-i\right); \\ i & \text{otherwise.} \end{cases} \tag{2.4}$$

Suppose now that the latter case holds, that is, $q_{0,0} < q_{1,0}$ and $q_{0,1} > q_{1,1}$. Then if $S_t(v) = i$,

$$S_{t+1}(v) = \begin{cases} 1-i & \text{if } n_t(v;i) > \lambda^{1-2i} n_t\left(v;1-i\right); \\ i & \text{otherwise.} \end{cases} \tag{2.5}$$

We refer to the system governed by the evolution rules from (2.4) as the *majority regime*; while we refer to the system described in (2.5) as the *minority regime.* We will tackle each of these systems separately. The intuition for these names arises from how each individual agent tends to think of its neighbours. In the majority regime, agents will tend to follow the strategies which are shared by the majority of their neighbours; however, in the minority regime agents will generally choose the least popular strategy seen across their neighbourhood. The value of λ determines the strength of this tendency. As previously mentioned in Section 1.1, we see that payoff matrices in the majority regime give rise to two Nash equilibria which are pure strategies: namely with both players playing simultaneously strategy 1 or strategy 0, thus they are co-ordination games. From this point of view, the parameter λ can be seen as some form of bias between the two pure Nash equilibria.

2.1.2 Degenerate Payoff Matrices

We briefly discuss the case of degenerate payoff matrices. We recall from the introduction that a payoff matrix Q is non-degenerate if one of the following hold:

(i) We have that $q_{0,0} > q_{1,0}$ and $q_{0,1} < q_{1,1}$.

(ii) Or, $q_{0,0} < q_{1,0}$ and $q_{0,1} > q_{1,1}$.

If neither condition holds then we say that Q is degenerate. The above classification leads to two possible degenerate matrices: either rows 0 and 1 are identical, or one of the rows contains both column maxima. Therefore if Q is degenerate, then the behaviour of the interacting node system is readily deduced. The system will either reach stability from the outset, or it will reach stable unanimity after one round. We summarise this behaviour in the following lemma.

Lemma 2.1.1. *Let Q be a degenerate payoff matrix, G a connected graph, and $\mathcal{S}$ an initial configuration of vertex strategies. Then the interacting node system $(G, Q, \mathcal{S})$, evolves as follows: Either the system is stable from $T = 0$, or the system is unanimous and stable from $T = 1$.*

Proof. We divide our proof into a number of cases. Firstly, we note that if $q_{0,0} = q_{1,0}$ and $q_{0,1} = q_{1,1}$, then we must achieve stability from the initial state, since $T_0(v) = T_0'(v)$. Hence, for all $v \in V(G)$ we have that $S_0(v) = S_t(v)$ for $t \geq 0$.

Suppose that the above case does not occur, and we have $q_{0,0} > q_{1,0}$. For Q to be degenerate, we are forced to have that $q_{0,1} \geq q_{1,1}$. Consequently, we now have top-row domination in Q. For any game played by vertex v, it is always optimal to play strategy zero. As G is connected, every vertex will play at least one game, and, in particular, it is always optimal to play strategy 0, which corresponds to the top row of the payoff matrix. Therefore, for all v we have that $S_1(v) = 0$, and a stable unanimity is achieved. A similar row domination argument follows for the remaining possibilities of degenerate matrices. □

As Lemma 2.1.1 holds when G is any connected graph, it then follows that this argument suffices as a proof of both Theorem 1.1.1 and Theorem 1.1.2 for the case that Q is a degenerate payoff matrix.

2.2 Skewed Interacting Node Systems

In this section, we are concerned with node systems where $\lambda \neq 1$. The central focus of this analysis is finding which sub-structures can block unanimity (i.e they evolve independently of the behaviour of the rest of the graph). We show that if none of these blocking structures are present, then the node system achieves unanimity. The emergence of these structures is characterised by Theorems 1.1.3 and 1.1.4. In this chapter, we will show that in the majority and minority regimes, there exists respective thresholds for the existence of these blocking structures. We remark that in each regime these structures, and hence their respective thresholds, are different. Furthermore, we then show that if these structures are not present, then the node system achieves unanimity within $\beta \log n$ rounds, as given by Theorem 1.1.2. This is done by first showing that a sub-linear majority is achieved in the first round, such a result hinges on the fact that $\lambda \neq 1$. Following from this, we partition the network into vertices of high and low degree, defined in terms of a suitable constant. We show that between each round, the size of the majority strategy among high-degree vertices increases until unanimity is achieved in the high-degree part, this is shown in Theorem 2.2.16. Following this, we show that given these blocking structures do not exist, the remaining low-degree vertices also synchronise with the high-degree vertices.

2.2.1 Blocking Structures and their Distribution

We will start with the identification of those structures/induced subgraphs of $G(n,p)$ which, roughly speaking, will stay immune to what the rest of the graph is doing. Thus, these substructures act as obstructions to unanimity.

In particular, the structure we consider is an (ℓ, k)-*blocking star*. This is a star whose central vertex has degree $\ell + k$ in $G(n, d/n)$ and, furthermore, ℓ leaves of the star have degree 1 inside $G(n, d/n)$, whereas we impose no restriction on the degrees of the remaining k leaves. We call the latter leaves the *connectors* of the blocking star, whereas the ℓ leaves of degree 1 are called the *blocking leaves*. Such a structure is illustrated in Figure 2.1.

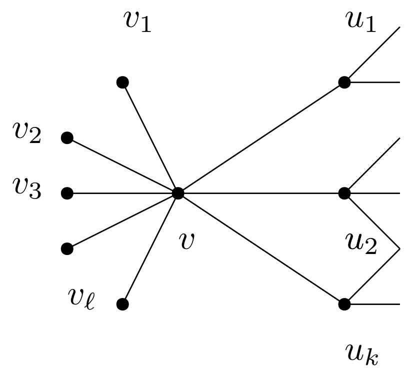

Figure 2.1: A blocking star with central vertex v, blocking leaves $v_1, \ldots, v_\ell$ and connectors $u_1, \ldots, u_k$.

An (ℓ, k)-blocking star has the property that it can block or absorb the influence of the external vertices depending on the choice of its parameters ℓ and k, respectively. Let us consider first the minority regime. It will turn out that it will be sufficient to consider $(1, k)$-blocking stars for a suitable defined k. Let $i^* \in \{0, 1\}$ be such that $\lambda^{2i^*-1} = \max\{\lambda, \lambda^{-1}\}$. For $i, j \in \{0, 1\}$, we say that a star has the (i, j)-*configuration* if the blocking leaves play strategy i whereas the centre plays strategy j.

Claim 2.2.1. *Consider a* $(1, k)$*-blocking star with* $k \leq \lfloor \lambda^{2i^*-1} \rfloor$*. In the minority regime, if the star ever gets into the* $(i^*, 1 - i^*)$*-configuration, it will stay so forever.*

Proof. Let v_1 denote the blocking leaf and v denote the centre of the star. Suppose that $S_t(v_1) = i^*$ but $S_t(v) = 1 - i^*$. By (2.5), the vertex v will change strategy, if $n_t(v; 1-i^*) > \lambda^{2i^*-1} n_t(v; i^*)$. Since $S_t(v_1) = i^*$, then $n_t(v; 1-i^*) \leq k \leq \lfloor \lambda^{2i^*-1} \rfloor$. But also $n_t(v; i^*) \geq 1$. So we would have $\lfloor \lambda^{2i^*-1} \rfloor > \lambda^{2i^*-1}$, which is impossible. Thus, $S_{t+1}(v) = 1 - i^*$. Furthermore, $S_{t+1}(v_1) = i^*$, since v_1 has no neighbours playing i^*. □

Claim 2.2.2. *Consider a $(1,k)$-blocking star with $k > \lfloor \lambda^{2i^*-1} \rfloor$. In the minority regime, if the k connectors simultaneously alternate between $1 - i^*$ and i^*, then the blocking leaf and centre will eventually synchronise with them.*

Proof. Let v_1 be the blocking leaf, v be the centre and $u_1, \ldots, u_k$ be k connectors of the star. Suppose that $S_t(u_j) = 1 - i^*$, for all $j = 1, \ldots, k$.

If $S_t(v) = 1 - i^*$, then v will change strategy if $n_t(v; 1-i^*) > \lambda^{2i^*-1} n_t(v; i^*)$ (cf. (2.5)). But $n_t(v; 1-i^*) \geq k$ and $n_t(v; i^*) \leq 1$. Since $k > \lfloor \lambda^{2i^*-1} \rfloor$, the above inequality is indeed satisfied. Hence, $S_{t+1}(v) = i^*$. Also, note that $S_{t+1}(v_1) = i^*$.

On the other hand, if $S_t(v) = i^*$, then v will not change strategy if $n_t(v; i^*) \leq \lambda^{1-2i^*} n_t(v; 1-i^*)$. But $n_t(v; 1-i^*) \geq k > \lfloor \lambda^{2i^*-1} \rfloor$. Therefore, $n_t(v; 1-i^*) > \lambda^{2i^*-1}$. So $\lambda^{1-2i^*} n_t(v; 1-i^*) > 1$. But $n_t(v; i^*) \leq 1$ and the inequality is satisfied. Therefore, $S_{t+1}(v) = i^*$. However, now $S_{t+1}(v_1) = 1 - i^*$.

Suppose now that $S_t(u_j) = i^*$, for all $j = 1, \ldots, k$. If $S_t(v) = 1 - i^*$, then by (2.5) v will change strategy if $n_t(v; 1-i^*) > \lambda^{2i^*-1} n_t(v; i^*)$. But now $n_t(v; 1-i^*) \leq 1$ and $n_t(v; i^*) \geq k > \lambda^{2i^*-1}$. Thus, the above inequality is not satisfied and $S_{t+1}(v) = 1 - i^*$. Also, note that $S_{t+1}(v_1) = i^*$.

If $S_t(v) = i^*$, then v will not change strategy if $n_t(v; i^*) \leq \lambda^{1-2i^*} n_t(v; 1-i^*)$. Now, $n_t(v; i^*) \geq k > \lambda^{2i^*-1} > 1$ but $\lambda^{1-2i^*} n_t(v; 1-i^*) \leq \lambda^{1-2i^*} < 1$. So the above inequality is not satisfied and $S_{t+1}(v) = 1 - i^*$. Furthermore, $S_{t+1}(v_1) = 1 - i^*$.

We thus conclude that in any case, the centre v will synchronise with the k connectors. Note that from the above four cases, we see that the blocking leaf v_1 will synchronise with v at the steps where it changes state. Hence, it will also synchronise with the k connectors too. □

In the majority regime, it will turn out that we will need to consider $(\ell, 1)$-blocking stars. We will show that if ℓ is sufficiently large, then the $(\ell, 1)$-blocking star can block the influence of the external vertices and retain the strategy of its vertices.

Claim 2.2.3. *Consider an $(\ell, 1)$-blocking star with $\ell \geq \lceil \lambda^{2i^*-1} \rceil$. In the majority regime, if it is set to the $(1-i^*, 1-i^*)$-configuration initially, then it will stay in this configuration forever.*

Proof. Suppose that $v_1, \ldots, v_\ell$ are the ℓ-blocking leaves and v is the centre of the star. Assume that all $S_t(v) = S_t(v_1) = \cdots = S_t(v_\ell) = 1 - i^*$. By (2.4), the centre will change strategy at step $t + 1$, if $n_t(v; 1 - i^*) < \lambda^{2i^*-1} n_t(v; i^*)$. But $n_t(v; 1 - i^*) = \ell \geq \lceil \lambda^{2i^*-1} \rceil$ and $n_t(v; i^*) \leq 1$. Hence, we should have $\lceil \lambda^{2i^*-1} \rceil < \lambda^{2i^*-1}$, which is impossible. So $S_{t+1}(v) = 1 - i^*$. Further, note that the ℓ-blocking leaves will adopt the strategy of the centre at step $t + 1$ since they have no other neighbours. Thus, $S_{t+1}(v_j) = 1 - i^*$, for any $j = 1, \ldots, \ell$, as well. □

Claim 2.2.4. *Consider an $(\ell, 1)$-blocking star with $\ell \geq 1$. In the majority regime, if it is set to the (i^*, i^*)-configuration initially, then it will stay in this configuration forever.*

Proof. Suppose that $v_1, \ldots, v_\ell$ are the ℓ-blocking leaves and v is the centre of the star. Assume that all $S_t(v) = S_t(v_1) = \cdots = S_t(v_\ell) = i^*$. By (2.4), the centre will change strategy at step $t+1$, if $n_t(v; i^*) < \lambda^{1-2i^*} n_t(v; 1-i^*)$. But $n_t(v; i^*) \geq 1$ and $n_t(v; 1-i^*) \leq 1$. So the above inequality is not satisfied, since $\lambda^{1-2i^*} < 1$, and, therefore, $S_{t+1}(v) = i^*$. Finally, the ℓ-blocking leaves will retain the strategy of the centre at step $t+1$, since they have no other neighbours. Thus, $S_{t+1}(v_j) = i^*$, for any $j = 1, \ldots, \ell$, as well. □

Now, we shall give a general condition that determines the distribution of the (ℓ, k)-blocking stars in $G(n, d/n)$. The following results will be useful both for the subcritical and the supercritical regime that we analyse in the next subsection.

Let $X_{\ell,k,n}$ be the random variable which is the number of (ℓ, k)-blocking stars in $G(n, d/n)$ and let $X^{(1)}_{\ell,k,n}$ be the number of those (ℓ, k)-blocking stars which are subgraphs of $L_1(G(n,d/n))$. Clearly, $X^{(1)}_{\ell,k,n} \le X_{\ell,k,n}$. However, we will show that a.a.s. these two random variables are approximately equal.

In particular, we will show the following three lemmas. The first Lemma provides an asymptotic for the expected number of (ℓ,k)-blocking stars in $G(n,d/n)$.

Lemma 2.2.5. *Let $\ell, k \in \mathbb{N}$ and $p = d/n$, where $1 \ll d = d(n) = O(\log n)$. Then we have,*

$$\mathbb{E}\left[X_{\ell,k,n}\right] \sim n\frac{d^{\ell+k}}{\ell!k!}e^{-d(\ell+1)}.$$

Our second Lemma describes the sharp threshold for the existence of an (ℓ,k)-blocking star, in terms of ℓ, k and n, inside of $G(n,d/n)$. We show that below this threshold an (ℓ,k)-blocking star will exist a.a.s. While above this threshold, no such stars will exist a.a.s. In the case that d is within a constant of the threshold, then we show the number of (ℓ,k)-blocking stars converges in distribution to a Poisson distribution by applying Theorem 1.4.2.

Lemma 2.2.6. *Furthermore, if $d = \frac{1}{\ell+1}\log n + \frac{\ell+k}{\ell+1}\log\log n + \omega(n)$, then the following hold.*

i. If $\omega(n) \to -\infty$ as $n \to +\infty$, then $\mathbb{E}\left[X_{\ell,k,n}\right] \to +\infty$ as $n \to +\infty$ and moreover, a.a.s. $X_{\ell,k,n} \ge \mathbb{E}\left[X_{\ell,k,n}\right]/2$.

ii. If $\omega(n) \to c \in \mathbb{R}$ as $n \to +\infty$, then $\mathbb{E}\left[X_{\ell,k,n}\right] \to \frac{e^{(\ell+1)c}}{\ell!k!}$ as $n \to \infty$ and

$$X_{\ell,k,n} \xrightarrow{d} \mathrm{Po}\left(\frac{e^{(\ell+1)c}}{\ell!k!}\right).$$

iii. If $\omega(n) \to +\infty$ as $n \to +\infty$, then $\mathbb{P}(X_{\ell,k,n} > 0) < 2e^{-(\ell+1)\omega(n)}$, for any n sufficiently large. Thus, a.a.s. $X_{\ell,k,n} = 0$.

We denote $X^{(2)}_{\ell,k,n} = X_{\ell,k,n} - X^{(1)}_{\ell,k,n}$. This quantity describes the number of (ℓ,k)-blocking stars found outside of the giant component. The following Lemma states that in expectation there are asymptotically much fewer (ℓ,k)-blocking stars in the rest of the graph.

Lemma 2.2.7. *We have* $\mathbb{E}\left[X^{(2)}_{\ell,k,n}\right] = o(\mathbb{E}\left[X_{\ell,k,n}\right])$.

Hence by applying Markov's inequality we have that a.a.s. $X^{(2)}_{\ell,k,n} \leq \mathbb{E}\left[X_{\ell,k,n}\right]/4$. By Lemma 2.2.5, if $d = \frac{1}{\ell+1}\log n + \frac{\ell+k}{\ell+1}\log\log n + \omega(n)$, with $\omega(n) \to -\infty$, then a.a.s.

$$X^{(1)}_{\ell,k,n} \geq \frac{1}{4}\mathbb{E}\left[X_{\ell,k,n}\right]. \tag{2.6}$$

Furthermore, note that any two (ℓ,k)-blocking stars can share only their connector vertices. So, if we consider the initial assignment of strategies, each (ℓ,k)-blocking star inside $L_1(G(n,d/n))$ will be set into (i,j)-configuration with probability $1/2^{\ell+1}$, independently of each other. Thus, the weak law of large numbers together with (2.6) implies the following.

Lemma 2.2.8. *Let* $p = d/n$, *where* $d = \frac{1}{\ell+1}\log n + \frac{\ell+k}{\ell+1}\log\log n + \omega(n)$, *for* $k,\ell \in \mathbb{N}$. *Let* $i,j \in \{0,1\}$.

i. If $\omega(n) \to -\infty$ *as* $n \to +\infty$, *then a.a.s. at least* $\frac{1}{8}\cdot\frac{1}{2^{\ell+1}}\mathbb{E}\left[X_{\ell,k,n}\right]$ *of the* (ℓ_λ,k)*-blocking stars inside* $L_1(G(n,d/n))$ *will be set into* (i,j)*-configuration at the beginning of the process.*

ii. If $\omega(n) \to c \in \mathbb{R}$ *as* $n \to +\infty$, *then the number of* (ℓ,k)*-blocking stars inside* $L_1(G(n,d/n))$ *will be set into* (i,j)*-configuration at the beginning of the process converges in distribution as* $n \to \infty$ *to a random variable distributed as*

$$\mathrm{Bin}\left(\mathrm{Po}(e^{(\ell+1)c}/(\ell!k!)), \frac{1}{2^{\ell+1}}\right).$$

We conclude this section with the proofs of Lemmas 2.2.5 and 2.2.7.

Proof of Lemma 2.2.5. We start with the expected value of $X_{\ell,k,n}$. Suppose S denotes a set of size $\ell+k+1$ on which an (ℓ,k)-blocking star will be formed. There are $\binom{n}{\ell+k+1}$ ways to select these vertices and $(\ell+k+1)\binom{\ell+k}{k}$ ways to select the centre v and the connector vertices $u_1, \ldots, u_k$. Suppose that the remaining vertices are $v_1, \ldots, v_\ell$. We have $\mathbb{P}(d(v_1) = 1 | v_1 \sim v) = (1-p)^{n-1}$. For $j = 2, \ldots, \ell$,

$$\mathbb{P}(d(v_j) = 1 | v_i \sim v, d(v_i) = 1, \text{for } i = 1, \ldots, j-1) = (1-p)^{n-(j-1)} < e^{-d+d\ell/n}.$$

But since $j \leq \ell$, we have for n sufficiently large

$$e^{-d-(d/n)^2} \leq (1-p)^{n-(j-1)} < e^{-d+d\ell/n}.$$

Hence, using the assumption that $d = O(\log n)$

$$\mathbb{P}(d(v_1) = \cdots = d(v_\ell) = 1 | v_i \sim v, \text{for } i = 1, \ldots, \ell) \sim e^{-d\ell}.$$

Also, $\mathbb{P}(d_{V_n \setminus S}(v) = 0) = (1-p)^{n-(\ell+k+1)} \sim e^{-d}$ Thus, we obtain

$$\begin{aligned}
\mathbb{E}\left[X_{\ell,k,n}\right] &\sim \frac{n^{\ell+k+1}}{(\ell+k+1)!} \cdot (\ell+k+1) \binom{\ell+k}{k} \cdot \left(\frac{d}{n}\right)^{\ell+k} \cdot e^{-d(\ell+1)} \\
&= \frac{n^{\ell+k+1}}{(\ell+k+1)!} \cdot (\ell+k+1) \frac{(\ell+k)!}{\ell! k!} \cdot \left(\frac{d}{n}\right)^{\ell+k} \cdot e^{-d(\ell+1)} \\
&= n \frac{d^{\ell+k}}{\ell! k!} \cdot e^{-d(\ell+1)}.
\end{aligned}$$

This concludes the proof of this lemma. □

Proof Of Lemma 2.2.6. Now, the value of $\lim_{n\to\infty} \mathbb{E}\left[X_{\ell,k,n}\right]$ is deduced as in parts $i., ii.$ and $iii.$ of the lemma follows by taking $d(\ell+1) = \log n + (\ell+k)\log\log n + (\ell+1)\omega(n)$, where either $\omega(n) \to +\infty$ or $\to c$ or $\to -\infty$, as $n \to \infty$, respectively.

For Part $iii.$, Markov's inequality implies that

$$\mathbb{P}(X_{\ell,k,n} > 0) \leq \mathbb{E}\left[X_{\ell,k,n}\right] \overset{n\ large}{<} 2 \cdot e^{-(\ell+1)\omega(n)} \to 0, \text{ as } n \to +\infty.$$

For Parts $i.$ and $ii.$ we now show that for any fixed integer $r \geq 2$, we have

$$\mathbb{E}\left[X_{\ell,k,n}(X_{\ell,k,n}-1)\cdots(X_{\ell,k,n}-(r-1))\right] \sim \mathbb{E}^r\left[X_{\ell,k,n}\right],$$

where $\mathbb{E}^r\left[X_{\ell,k,n}\right]$ is defined as $\left(\mathbb{E}\left[X_{\ell,k,n}\right]\right)^r$. So the second statement in Part $i.$ will follow from Chebyshev's inequality as for $r=2$, the above implies that $\mathrm{Var}(X_{\ell,k,n}) = o(\mathbb{E}\left[X_{\ell,k,n}\right])$. The second statement in Part $ii.$ will follow from Theorem 1.4.2.

Consider r subsets $S_1,\ldots,S_r \subset V_n$ of size $\ell+k+1$. They may all induce (ℓ,k)-blocking stars only if any two of them share at most k vertices. For $S \subset V_n$ let I_S be the indicator random variable that is equal to 1 if and only if S is an (ℓ,k)-blocking star. If $S_i \cap S_j = \varnothing$ for all $i \neq j$, then

$$\mathbb{P}(I_{S_1}=\cdots=I_{S_r}=1) \sim \left((\ell+k+1)\binom{\ell+k}{k}p^{(\ell+k)}e^{-d(\ell+1)}\right)^r. \tag{2.7}$$

There are $\prod_{i=1}^r \binom{n-(i-1)(\ell+k+1)}{\ell+k+1}$ to select the ordered r-tuple of pairwise disjoint sets $(S_1,\ldots,S_r)$. But

$$\prod_{i=1}^r \binom{n-(i-1)(\ell+k+1)}{\ell+k+1} \sim \left(\frac{n^{\ell+k+1}}{(\ell+k+1)!}\right)^r. \tag{2.8}$$

Now, let us assume that $S_i \cap S_j \neq \varnothing$ for some $i \neq j$. Note first that if $|S_i \cap S_j| > k$, then the two sets cannot induce (ℓ,k)-blocking stars simultaneously. If $S_i \cap S_j = \{u_1,\ldots,u_s\}$, with $s \leq k$ and S_i, S_j induce (ℓ,k)-blocking stars, then $u_1,\ldots,u_s$ must be connector vertices in both of them. Since ℓ and k are fixed, there are $O(1)$ ways to select the vertices in $\cup_{i=1}^r S_i$ that will be the connectors of the (ℓ,k)-blocking stars. Let $S' \subset \cup_{i=1}^r S_i$ be such a choice; by the above observation, this contains any vertex which belongs to at least two members of the r-tuple. Set $S^{1..r} = \cup_{i=1}^r S_i$. Hence,

$$\mathbb{P}(I_{S_1}=\cdots=I_{S_r}=1) = O(1)\cdot\left[(1-p)^{\binom{\ell+k+1}{2}-(\ell+k)}p^{(\ell+k)}\right]^r \cdot \mathbb{P}(\forall v \in S^{1..r}\backslash S', d_{V_n\backslash S^{1..r}}(v)=0).$$

Note that $|S^{1..r} \setminus S'| = r(\ell+1)$. So, the latter probability is

$$\mathbb{P}(\forall v \in S^{1..r} \setminus S', d_{V_n \setminus S^{1..r}}(v) = 0) = (1-p)^{r(\ell+1)(n-|S^{1..r}|)} \sim e^{-dr(\ell+1)}.$$

Therefore,

$$\mathbb{P}(I_{S_1} = \cdots = I_{S_r} = 1) = O(1) \cdot p^{r(\ell+k)} e^{-dr(\ell+1)}. \tag{2.9}$$

Now, such an ordered r-tuple can be selected into at most

$$\binom{n}{r(\ell+k+1)-1} = o(n^{r(\ell+k+1)}) \tag{2.10}$$

ways. Thus,

$$\begin{aligned}
\mathbb{E}(X_{\ell,k,n}(X_{\ell,k,n}-1)\cdots(X_{\ell,k,n}-(r-1))) &= \sum_{(S_1,\ldots,S_r): S_i \subset V_n, |S_i| = \ell+k+1} \mathbb{P}(I_{S_1} = \cdots = I_{S_r} = 1) \\
&= (1+o(1)) \left(\frac{n^{\ell+k+1}}{(\ell+k+1)!} \cdot \left((\ell+k+1) \binom{\ell+k}{k} \right) p^{\ell+k} e^{-d(\ell+1)} \right)^r \text{ by (2.7),(2.8)} \\
&\quad + o(n^{r(\ell+k+1)}) p^{r(\ell+k)} e^{-dr(\ell+1)} \text{ by (2.9),(2.10)} \\
&\sim \mathbb{E}^r(X_{\ell,k,n}).
\end{aligned}$$

□

Proof of Lemma 2.2.7. In this lemma, we will bound the expected number of (ℓ, k)-blocking stars which do not belong to $L_1(G(n, d/n))$. Recall that the random variable which counts these is $X^{(2)}_{\ell,k,n}$.

We will give an upper bound on the expected number of connected components of order at most $\log n$ which contain an (ℓ, k)-blocking star. This suffices due to the following result, (Theorem 6.10 in [15]) about the structure of $G(n, d/n)$ in the super-critical regime.

Theorem 2.2.9. *Let $p = d/n$ with $d \gg 1$. Then a.a.s. all connected components of $G(n, d/n)$ apart from $L_1(G(n, d/n))$ have order at most $\log n$.*

For $r \geq \ell+k+1$, let $C_{\ell,k,r}$ denote the number of connected components which contain an (ℓ, k)-blocking star and have order r. We will give an upper bound on the expected value of $C_{\ell,k,r}$. For two sets $S \subset S' \subset V_n$ having $|S| = \ell + k + 1$ and $|S'| = r$, and $v, u_1, \ldots, u_k \in S$, we set $I(S, S', v, u_1, \ldots, u_k)$ to be the indicator random variable which is equal to 1 if and only if S' is a connected component in $G(n, d/n)$, where in particular, S induces an (ℓ, k)-blocking star with centre v and connectors $u_1, \ldots, u_k$. Let $\mathcal{S}_{\ell,k,r}$ denote the set of pairs (S, S') such that $S \subset S' \subset V_n$ with $|S| = \ell+k+1$ and $|S'| = r$. Moreover, for a set $S \subset V_n$ of size at least $k+1$, let $S_{\neq}^{(k+1)}$ denote the set of all ordered $k+1$-tuples of distinct vertices in S. With this notation, we can write

$$\mathbb{E}\left[C_{\ell,k}\right] \leq \sum_{(S,S') \in \mathcal{S}_{\ell,k,r}} \sum_{(v,u_1,\ldots,u_k) \in S_{\neq}^{(k+1)}} \mathbb{E}\left[I(S', S, v, u_1, \ldots, u_k)\right]. \tag{2.11}$$

For $S \subset V_n$ having $|S| = \ell+k+1$ and $(v, u_1, \ldots, u_k) \in S_{\neq}^{(k+1)}$, we let $I_1(S, v, u_1, \ldots, u_k)$ be the indicator random variable that is equal to 1 if and only if S forms an ℓ-blocking star with centre v and connectors $u_1, \ldots, u_k$. Also, we take $I_2(S')$ to be the indicator random variable that is equal to 1 if and only if S' is a connected component of $G(n, d/n)$.

For a given $(S, S') \in \mathcal{S}_{\ell,k}$ and $(v, u_1, \ldots, u_k) \in S_{\neq}^{(k+1)}$, we write

$$\begin{aligned}
&\mathbb{P}(I(S', S, v, u_1, \ldots, u_k) = 1) = \\
&\qquad \mathbb{P}(I_1(S, v, u_1, \ldots, u_k) = 1) \cdot \mathbb{P}(I_2(S') = 1 \mid I_1(S, v, u_1, \ldots, u_k) = 1).
\end{aligned} \tag{2.12}$$

We will provide an upper bound on $\mathbb{P}(I_2(S') = 1 \mid I_1(S, v, u_1, \ldots, u_k) = 1)$. If S induces an (ℓ, k)-blocking star with centre v and connectors $u_1, \ldots, u_k$, then any spanning tree $T_{S'}$ on S' contains the star on $S \setminus \{u_1, \ldots, u_k\}$ centred at v as an induced subgraph and. Moreover, one of the edges $u_i v$ is a cutting edge between $S' \setminus (S \setminus \{u_1, \ldots, u_k\})$ and $S \setminus \{u_1, \ldots, u_k\}$. Thus, $T_{S'} \setminus (S \setminus \{u_1, \ldots, u_k\})$ is a spanning tree of the subgraph induced by the set $S' \setminus (S \setminus \{u_1, \ldots, u_k\})$. Furthermore, if S' is a connected component in $G(n, d/n)$, there are no edges between $S' \setminus (S \setminus \{u_1, \ldots, u_k\})$ and $V_n \setminus S'$.

With the above observations and $p = d/n$ we can give the following bound:

$$\mathbb{P}(I_2(S') = 1 \mid I_1(S, v, u_1, \ldots, u_k) = 1) \leq (r - (\ell+1))^{r-(\ell+1)-2} p^{r-(\ell+1)-1} (1-p)^{(n-r)(r-(\ell+1))}, \quad (2.13)$$

since there are k^{k-2} labelled trees on k vertices. We have that

$$(r - (\ell+1))^{r-(\ell+1)-2} p^{r-(\ell+1)-1} \leq (rp)^{r-(\ell+1)-1} = (rp)^{r-\ell-2},$$

and for $r \leq \log n$,

$$(1-p)^{(n-r)(r-(\ell+1))} \sim e^{-dr} \leq 2e^{-d(r-\ell-1)} = 2e^{-d(r-\ell-2)} e^{-d}.$$

Using these in (2.13) we get

$$\mathbb{P}(I_2(S') = 1 \mid I_1(S, v, u_1, \ldots, u_k) = 1) \leq 2(rpe^{-d})^{r-\ell-2} e^{-d}.$$

Hence, the left-hand side in (2.12) is bounded as

$$\mathbb{P}(I(S', S, v, u_1, \ldots, u_k) = 1) \leq \mathbb{P}(I_1(S, v, u_1, \ldots, u_k) = 1) \cdot \left(2(rpe^{-d})^{r-\ell-2} e^{-d}\right).$$

Thus, for n sufficiently large, (2.11) yields:

$$\begin{aligned} \mathbb{E}(C_{\ell,k,r}) &\leq \sum_{(S,S') \in \mathcal{S}_{\ell,k,r}} \sum_{(v,u_1,\ldots,u_k) \in S_{\neq}^{(k+1)}} \mathbb{P}(I_1(S, v, u_1, \ldots, u_k) = 1) \cdot \left(2(rpe^{-d})^{r-\ell-2} e^{-d}\right) \\ &= \sum_{S \subset V_n : |S| = \ell+k+1} \sum_{(v,u_1,\ldots,u_k) \in S_{\neq}^{(k+1)}} \mathbb{P}(I_1(S, v, u_1, \ldots, u_k) = 1) \times \\ &\qquad \sum_{S' : |S'| = r, S \subset S'} 2(rpe^{-d})^{r-\ell-2} e^{-d}. \end{aligned} \quad (2.14)$$

For a fixed choice of S, we will provide an upper bound on the inner sum. This is

$$\begin{aligned}\sum_{S':|S'|=r,S\subset S'}(rpe^{-d})^{r-\ell-2}e^{-d} &\leq \sum_{r=\ell+k+1}^{\log n}\binom{n-\ell-k-1}{r-\ell-k-1}(rpe^{-d})^{r-\ell-k-1}e^{-d}\\ &\leq e^{-d}\sum_{r=\ell+k+2}^{\log n}\left(\frac{ne}{r-\ell-k-1}\right)^{r-\ell-k-1}\left(\frac{rde^{-d}}{n}\right)^{r-\ell-k-1}\\ &\quad + e^{-d}((\ell+k+1)pe^{-d}).\end{aligned}$$

By observing that $d \gg 1$ we have that the second term is $o(1)$. We also observe that $\frac{r}{r-\ell-k-1} \leq \ell+k+2$ hence by shifting the index we have that,

$$\sum_{S':|S'|=r,S\subset S'}(rpe^{-d})^{r-\ell-2}e^{-d} \leq e^{-d}\sum_{j=1}^{\log n+\ell+k+1}\left((\ell+k+2)ede^{-d}\right)^j + o(1).$$

Furthermore by observing that $(\ell+k+2)ede^{-d} \leq e^{-d/2}$ and bounding from above by a geometric series, we have

$$e^{-d}\sum_{j=1}^{\log n+\ell+k+1}\left((\ell+k+2)ede^{-d}\right)^j + o(1) \leq e^{-d}\sum_{j=1}^{\log n+\ell+2}\left(e^{-d/2}\right)^j + o(1) = o(1).$$

Note that this upper bound holds for n sufficiently large uniformly over all choices of S. Thus,

$$\mathbb{E}\left[C_{\ell,k,r}\right] = o(1)\cdot\sum_{S\subset V_n:|S|=\ell+k+1}\ \sum_{(v,u_1,\ldots,u_k)\in S_{\neq}^{(k+1)}}\mathbb{P}(I_1(S,v,u_1,\ldots,u_k)=1) = o(1)\cdot\mathbb{E}\left[X_{\ell,k,n}\right],$$

which concludes the proof of the lemma. $\square$

2.2.2 Small Degree Vertices, their Structure and their Role in Unanimity

In this subsection, we proceed with the proof of Theorems 1.1.2, 1.1.3 and 1.1.4. Suppose that $\mathcal{I} = (G, Q, \mathcal{S})$ is an interacting node system with $\lambda = \lambda(Q) \neq 1$ and $G = G(n,p)$.

For each $v \in V_n$ and for $\delta \in (0,1)$ we say that v is δ-*balanced* if for all $i \in \{0,1\}$, we have that $\big|n_0(v;i) - \mathbb{E}[n_0(v;i)]\big| \leq \delta\mathbb{E}[n_0(v;i)]$. If v is not δ-balanced then we say that v is δ-*unbalanced*, and we denote the set of δ-unbalanced vertices as $\mathcal{U}_\delta$. We denote P_t and N_t to denote the subsets of vertices whose agents play strategy 1 and 0, respectively, at the t^{th} step. More formally, for $t \geq 0$, we define $P_t = \{v \in V_n : S_t(v) = 1\}$ and $N_t = \{v \in V_n : S_t(v) = 0\}$. We define $m_t = N_t$ if $|N_t| \leq |P_t|$, and P_t otherwise. Hence m_t is the set of vertices playing the minority strategy at time t, or the vertices playing zero in the case $|N_t| = |P_t|$. We define $\mu_t = \min\{|N_t|, |P_t|\}$. The following lemma describes the first round of evolution; notably, it describes the formation of a large majority after a single round. Recall that $i^* \in \{0,1\}$ is the strategy which satisfies $\lambda^{1-2i^*} < 1$. Note that since $\lambda \neq 1$, exactly one of 0 or 1 satisfies this. Note further that $\lambda^{1-2i^*} = \min\{\lambda, \lambda^{-1}\}$. Hence,

$$\max\{\lambda, \lambda^{-1}\} \cdot \lambda^{1-2i^*} = \max\{\lambda, \lambda^{-1}\} \cdot \min\{\lambda, \lambda^{-1}\} = \lambda \cdot \lambda^{-1} = 1. \tag{2.15}$$

We will use this identity later on. Now, we will show that after one round, in the majority regime strategy, i^* will become the dominant strategy among the vertices of $G(n, d/n)$. However, in the minority regime, it will be strategy $1 - i^*$ that will dominate.

Lemma 2.2.10. *Let $p = d/n$ where $d \gg 1$. For any $0 < \lambda \neq 1$, there exists $\gamma > 0$ for which the following holds. A.a.s. across the product space of $G(n,p)$ and $\mathcal{S}_{1/2}$: For the interacting node system $\mathcal{I} = (G(n,p), Q, \mathcal{S}_{1/2})$ with $\lambda(Q) = \lambda$, we have $\mu_1 \leq ne^{-\gamma d}$. In particular, after the first round, the majority of the vertices will be playing either strategy i^* (majority regime) or strategy $1-i^*$ (minority regime). Furthermore, there exists $\alpha(\lambda) > 1$ such that if $d > \alpha(\lambda)\log n$, then a.a.s. $\mu_1 = 0$.*

Proof. Suppose first that our system is in the majority regime. By (2.4), if a vertex plays strategy $1 - i^*$, then it will change strategy if $n_0(v; 1 - i^*) < \lambda^{1-2(1-i^*)} n_0(v; i^*) = \lambda^{2i^*-1} n_0(v; i^*)$. Also, if a vertex plays strategy i^*, then it will stay there, if $n_0(v; i^*) \geq \lambda^{1-2i^*} n_0(v; 1 - i^*)$.

Suppose now that our system is in the minority regime. By (2.5), if a vertex plays strategy $1-i^*$, then it will remain there if $n_0(v;1-i^*) \leq \lambda^{1-2(1-i^*)} n_0(v;i^*) = \lambda^{2i^*-1} n_0(v;i^*)$. Also, if a vertex plays strategy i^*, then it will switch to strategy $1-i^*$, if $n_0(v;i^*) > \lambda^{1-2i^*} n_0(v;1-i^*)$. Now, note that if v is δ-balanced, then provided that $\delta = \delta(\lambda) \in (0,1)$ is sufficiently small

$$\lambda^{1-2i^*} < \frac{1-\delta}{1+\delta} \leq \frac{n_0(v;1-i^*)}{n_0(v;i^*)} \leq \frac{1+\delta}{1-\delta} < \lambda^{2i^*-1}.$$

In other words, if v is δ-balanced, then all the above four inequalities will be satisfied. We thus arrive at the following conclusions:

1. if v is δ-balanced and Q is in the majority regime, then $S_1(v) = i^*$;

2. if v is δ-balanced but Q is in the minority regime, then $S_1(v) = 1-i^*$.

Furthermore, if the majority of the vertices in V_n are δ-balanced, then $\mu_1 = |\mathcal{U}_\delta|$. We will show that a.a.s. the majority of the vertices in V_n are δ-balanced, whereby they adopt strategy i^* or $1-i^*$ after one step, as described above. We will show that an arbitrary vertex $v \in V_n$ is δ-balanced with probability $1-o(1)$. For any $v \in V_n$ the random variables $n_0(v;0)$ and $n_0(v;1)$ have identical distributions, namely the binomial distribution $\mathrm{Bin}(n-1, p/2)$. Therefore

$$\mathbb{E}[n_0(v;1)] = \mathbb{E}[n_0(v;0)] = (1-o(1))\frac{d}{2}.$$

Set $\bar{\gamma} = \delta^2/7$. We bound the probability that $v \in \mathcal{U}_\delta$. By Chernoff's inequality (1.3) we have:

$$\mathbb{P}\Big(\big|n_0(v;0) - \mathbb{E}[n_0(v;0)]\big| \geq \delta \mathbb{E}[n_0(v;0)]\Big) \leq 2e^{-\frac{\delta^2}{3}\frac{(1-o(1))d}{2}} \leq e^{-\bar{\gamma} d},$$

where the last inequality holds for n sufficiently large. The same holds for $n_0(v;1)$ as it is identically distributed to $n_0(v;0)$. Hence, by the union bound, for any $v \in V_n$

$$\mathbb{P}\left[v \in \mathcal{U}_\delta\right] \leq 2e^{-\bar{\gamma} d}.$$

Therefore, $\mathbb{E}[|\mathcal{U}_\delta|] \leq 2ne^{-\bar{\gamma}d}$. By Markov's inequality we have:

$$\mathbb{P}(|\mathcal{U}_\delta| \geq ne^{-\bar{\gamma}d/2}) \leq \frac{2ne^{-\bar{\gamma}d}}{ne^{-\bar{\gamma}d/2}} = 2e^{-\bar{\gamma}d/2} = o(1).$$

Therefore, a.a.s. we have that $|\mathcal{U}_\delta| < ne^{-\bar{\gamma}d/2}$. In turn, a.a.s. $\mu_1 \leq |\mathcal{U}_\delta| < ne^{-\bar{\gamma}d/2}$. Finally, note that the last inequality above implies that if $d > \alpha(\lambda)\log n$, for some $\alpha(\lambda) > 1$ sufficiently large then the above is $o(1)$, which shows the last part of the lemma. □

For the analysis of the subsequent rounds we will split the vertices of $G(n, d/n)$ into two classes and we will consider their evolution separately. More specifically, for $C \in \mathbb{N}$ we set

$$H_n(C) := H(C, G(n, d/n)) := \{v \in V_n \ : \ d(v) \geq C\},$$

and

$$L_n(C) := L(C, G(n, d/n)) := V_n \setminus H(C, G(n, d/n)) = \{v \in V_n \ : \ d(v) < C\}.$$

It is of note that a large proportion of the vertices will fall into $H_n(C)$.

Claim 2.2.11. *Let $G \sim G(n, d/n)$ for $d \gg 1$. Then for any positive constant C we have that a.a.s $|H_n(C)| \geq n(1 - o(1))$.*

Proof. Clearly we have that $|L_n(C)| + |H_n(C)| = n$, therefore it suffices to show that for any positive constant C, a.a.s $|L_n(C)| = o(n)$. Suppose $v \in V_n$, then by the Chernoff bound with $\delta = 1/2$, we have following:

$$\mathbb{P}[d(v) \leq C] \leq \mathbb{P}\left(d(v) \leq \frac{d + o(1)}{2}\right) \leq e^{-\frac{d+o(1)}{12}} \leq e^{-d/24}.$$

Therefore it follows that $\mathbb{E}[|L_n(C)|] \leq ne^{-d/24}$, hence for any $\varepsilon > 0$, we have by Markov's:

$$\mathbb{P}(|L_n(C)| > \varepsilon n) \leq \frac{ne^{-d/24}}{\varepsilon n} \overset{d \gg 1}{=} o(1).$$

Given that $d \gg 1$, for any positive C we have that a.a.s $|L_n(C)| = o(n)$, therefore $|H_n(C)| = n(1 - o(1))$.

□

By Lemma 2.2.10, it suffices to assume that $d \leq \alpha(\lambda) \log n$, as otherwise a.a.s. the process reaches unanimity after one step. We will now provide some lemmas regarding the structure of the subgraph induced by the vertices in $L_n(C)$, for any fixed integer $C \geq 2$. Before doing this we shall give a bound on the joint probability that a given collection of vertices $S \subset V_n$ of size $|S| = O(1)$ belong to $L_n(C)$.

Claim 2.2.12. *Let $S \subset V_n$ be such that $|S| < k$, for some fixed $k \in \mathbb{N}$. Then*

$$\mathbb{P}(\forall v \in S, d(v) \leq C) \leq \left(2C\,(de)^C \cdot e^{-d}\right)^{|S|}.$$

Proof. If a vertex in S has degree at most C, then it has also degree at most C in $V_n \setminus S$. So we can write:

$$\mathbb{P}(\forall v \in S, d(v) \leq C) \leq \mathbb{P}(\forall v \in S, d_{V_n \setminus S}(v) \leq C).$$

Observe that these degrees form an independent family as they are determined by mutually disjoint sets of edges. Thereby,

$$\mathbb{P}(\forall v \in S, d_{V_n \setminus S}(v) \leq C) = \prod_{v \in S} \mathbb{P}(d_{V_n \setminus S}(v) \leq C).$$

But $d_{V_n \setminus S}(v)$ is distributed as $\mathrm{Bin}(n - |S|, d/n)$ and, therefore, its expected value is $d - o(1)$. Hence, $\mathbb{P}(d_{V_n \setminus S}(v) = k)$ is increasing as a function of k, if $d/k \to \infty$, as $n \to \infty$. Using $\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k$ we can write the following bound:

$$\begin{aligned} \mathbb{P}(d_{V_n \setminus S}(v) \leq C) &\leq C \cdot \binom{n}{C} \left(\frac{d}{n}\right)^C \left(1 - \frac{d}{n}\right)^{n - |S| - C} \\ &\overset{|S|<k}{\leq} C \left(\frac{ne}{C}\right)^C \left(\frac{d}{n}\right)^C e^{-d + o(1)} \end{aligned}$$

$$\leq \quad 2C\left(\frac{de}{C}\right)^C \cdot e^{-d}, \tag{2.16}$$

for n sufficiently large. Therefore,

$$\mathbb{P}(\forall v \in S, d_{V_n\setminus S}(v) \leq C) \leq \left(2C\,(de)^C \cdot e^{-d}\right)^{|S|},$$

and the claim now follows. □

The following lemmas describe the structure of the subgraph induced by the vertices in $L_n(C)$. We recall the definitions,

$$\ell_\lambda := \lceil \max\{\lambda, \lambda^{-1}\}\rceil \text{ and } c_\lambda := \frac{1}{\ell_\lambda + 1}.$$

Lemma 2.2.13. *Suppose that $p = d/n$ with $c_\lambda \log n \leq d \leq \alpha(\lambda)\log n$. A.a.s. there are no $\ell_\lambda + 2$ vertices in $L_n(C)$ that have a common neighbour.*

Proof. We will use a first-moment argument to bound the expected number of collections of vertices in L_n of size $\ell_\lambda + 2$ that have a common neighbour. Let $S \subset V_n$ be a subset of vertices. We denote the degree of v outside the set S by $d_{V_n\setminus S}(v)$. Let $S \subset V_n$ be such that $S = \{u_1, \ldots, u_{\ell_\lambda+2}\}$ and $z \in V_n \setminus S$. The expected number of collections of $\ell_\lambda + 2$ vertices in $L_n(C)$ which have a common neighbour is at most

$$\begin{aligned}
&\binom{n}{\ell_\lambda + 2} \cdot (n - (\ell_\lambda + 2)) \cdot \left(\prod_{i=1}^{\ell_\lambda+2} \mathbb{P}(u_i \sim z)\right) \cdot \mathbb{P}\left(\forall i = 1, \ldots, \ell_{\lambda+2}\ d_{V_n}(u_i) \leq C\right), \\
&\overset{Claim\ 2.2.12}{\leq} \quad n^{\ell_\lambda+3} \cdot \left(\frac{d}{n}\right)^{\ell_\lambda+2} \left(2C\,(de)^C \cdot e^{-d}\right)^{\ell_\lambda+2}, \\
&= \quad nd^{\ell_\lambda+2} \cdot \left(2C\,(de)^C \cdot e^{-d}\right)^{\ell_\lambda+2} = n \cdot e^{-d(\ell_\lambda+2)+O(\log d)}. \qquad (2.17)
\end{aligned}$$

But $d \geq c_\lambda \log n = (\ell_\lambda + 1)^{-1}\log n$. So $d(\ell_\lambda + 2) - \log n = \Omega(\log n)$, whereby the above expected value is $o(1)$. □

The above lemma implies in particular that a.a.s. at most $\ell_\lambda + 1$ vertices in $L_n(C)$ are adjacent to each vertex in $H_n(C)$. Thus, we see that if all vertices in $H_n(C)$ play a certain strategy simultaneously, then if C is large compared to ℓ_λ, then they may stay unaffected by what the vertices in $L_n(C)$ do.

Lemma 2.2.14. *Let $\ell \in \mathbb{N}$ and let $p = d/n$ where $\frac{1}{\ell+1}\log n \leq d = d(n) \leq \alpha(\lambda)\log n$. A.a.s. all connected sets of vertices in $L_n(C)$ have size at most $\ell + 1$.*

Proof. We will show that a.a.s. there are no connected sets of vertices in $L_n(C)$ of size $\ell + 2$ or more. If there is such a set, then in fact there must also be such a set of size exactly $\ell + 2$. So it suffices to show that a.a.s. no such subsets exist.

Let $S \subset V_n$ have $|S| = \ell + 2$. Then

$$\mathbb{P}(S \text{ is connected and } \forall v \in S, d(v) \leq C) \leq \mathbb{P}(S \text{ is connected}) \cdot \mathbb{P}(\forall v \in S, d(v) \leq C),$$

by the FKG inequality (Theorem 1.4.1), since the graph property that $\{S \text{ is connected}\}$ is non-decreasing whereas the property that $\{\forall v \in S, d(v) \leq C\}$ is non-increasing. Now,

$$\mathbb{P}(S \text{ is connected}) \leq |S|^{|S|-2} \cdot \left(\frac{d}{n}\right)^{|S|-1},$$

since if S induces a connected subgraph, then this has to have a spanning tree (selected in $|S|^{|S|-2}$ ways). By Claim 2.2.12, we have

$$\mathbb{P}(\forall v \in S, d(v) \leq C) \leq \left(2C\,(de)^C \cdot e^{-d}\right)^{|S|}.$$

Therefore,

$$\mathbb{P}(S \text{ is connected and } \forall v \in S, d(v) \leq C) \leq (\ell+2)^{\ell} \cdot \left(\frac{d}{n}\right)^{\ell+1} \cdot \left(2C\,(de)^C \cdot e^{-d}\right)^{\ell+2}.$$

Hence, the expected number of such subsets is at most

$$O(1) \cdot \binom{n}{\ell+2} \cdot \left(\frac{d}{n}\right)^{\ell+1} \cdot \left(d^C \cdot e^{-d}\right)^{\ell+2} = ne^{-(\ell+2)d + O(\log\log n)}.$$

But $d \geq \frac{1}{\ell+1} \log n$. Thereby, $(\ell+2)d - \log n = \Omega(\log n)$, and the right-hand side is $o(1)$. $\square$

Lemma 2.2.15. *Let $p = d/n$ where $c_\lambda \log n \leq d = d(n) \leq \alpha(\lambda) \log n$. A.a.s. all connected sets of vertices in $L_n(C)$ induce trees.*

Proof. By the previous lemma, it suffices only to consider sets of size at most $\ell_\lambda + 1$. Let $S \subset V_n$ with $|S| \leq \ell_\lambda + 1$. Then

$$\mathbb{P}(e(S) \geq |S| \text{ and } \forall v \in S, d(v) \leq C) \leq \mathbb{P}(e(S) \geq |S|) \cdot \mathbb{P}(\forall v \in S, d(v) \leq C),$$

by the FKG inequality (Theorem 1.4.1), since the graph property that $\{e(S) \geq |S|\}$ is non-decreasing and the property that $\{\forall v \in S, d(v) \leq C\}$ is non-increasing. But

$$\mathbb{P}(e(S) \geq |S|) = O(1) \cdot \left(\frac{d}{n}\right)^{|S|}.$$

Combining this with Claim 2.2.12 we get

$$\mathbb{P}(e(S) \geq |S| \text{ and } \forall v \in S, d(v) \leq C) = O(1) \cdot \left(\frac{d}{n}\right)^{|S|} \left(d^C e^{-d}\right)^{|S|} = O(1) \cdot n^{-|S|} e^{-d|S| + O(\log \log n)}.$$

As there are $\binom{n}{|S|} \leq n^{|S|}$ choices for S, the expected number of such sets is

$$O(1) e^{-d|S| + O(\log \log n)} = o(1).$$

The lemma follows from the union bound, taking the union over all possible values of $|S| \leq \ell_\lambda + 1$. $\square$

The next lemma will help us deal with the evolution of the vertices in $H_n(C)$. It shows that if one of the strategies occupies only a sublinear number of vertices in $H_n(C)$ (and we have almost unanimity) then after one more round the size of the minority strategy will contain only a fraction of these vertices.

Lemma 2.2.16. *Let $d = d(n)$ be such that $1 \ll d \leq \alpha(\lambda) \log n$, and let Q be a 2×2 non-degenerate payoff matrix. For any $\varepsilon > 0$ there exists $C_{\varepsilon,\lambda} \in \mathbb{N}$ such that for any $\gamma > 0$*

and for any $C \geq C_{\varepsilon,\lambda}$ a.a.s. $G(n, d/n)$ satisfies the following: for any initial configuration $\mathcal{S}$ with $|m_0 \cap H_n(C)| < ne^{-\gamma d}$, the interacting node system $\mathcal{I} = (G(n, d/n), Q, \mathcal{S})$ will have $|m_1 \cap H_n(C)| \leq \varepsilon |m_0 \cap H_n(C)|$.

Proof of Lemma 2.2.16. Suppose that initially, the majority strategy is i. Assume first that $\mathcal{I}$ is in the majority regime. Then in the random graph $G(n, d/n)$ typically a vertex is expected to have many more neighbours among those playing strategy i than those playing strategy $1 - i$. So one would expect that most vertices will adopt strategy i in the next round. If we revisit (2.4), we will see that if this does not happen, then $n_0(v; i) \leq \lambda^{1-2i} n_0(v; 1)$. Indeed, if v initially was playing strategy i, it switches to $1 - i$, if $n_0(v; i) < \lambda^{1-2i} n_0(v; 1)$. If v was initially playing $1 - i$, then it does not switch if $n_0(v; 1 - i) \geq \lambda^{1-2(1-i)} n_0(v; i)$. Rearranging the latter, we also get that $n_0(v; i) \leq \lambda^{1-2i} n_0(v; 1 - i)$.

Suppose now that $\mathcal{I}$ is in the minority regime. In this case, one would expect that most vertices will adopt strategy $1 - i$. Suppose that a vertex v initially plays i. By (2.5), it keeps on playing i after one round, if $n_0(v; i) \leq \lambda^{1-2i} n_0(v; i - 1)$. Similarly, if v initially plays $1 - i$, then it switches to i, if $n_0(v; 1 - i) > \lambda^{1-2(1-i)} n_0(v; i)$. If we rearrange the latter, we get $n_0(v; i) < \lambda^{1-2i} n_0(v; 1 - i)$. Furthermore, to reduce notation we set $H_n := H_n(C)$, where C is to be determined later. Also, we set $L_n := L_n(C)$.

Assuming that initially the most popular strategy in H_n is i, we say that a vertex $v \in V_n \cap H_n$ is *i-atypical*, if $n_0(v; i) \leq \lambda^{1-2i} n_0(v; 1 - i)$. Let $A_n^{(i)}$ denote the set of i-atypical vertices. We will show that a.a.s. for any assignment of strategies to H_n, respecting $|m_0| < ne^{-\gamma d}$, we have that $|A^{(i)}| \leq \varepsilon |m_0 \cap H_n|$. If this happens, then all but at most $\varepsilon |m_0 \cap H_n|$ vertices in H_n will behave as expected and, therefore, $|m_1 \cap H_n| \leq \varepsilon |m_0 \cap H_n|$.

We proceed with showing the above. We assume the majority strategy initially is i. So a vertex is i-atypical if $n_0(v; i) \leq \lambda^{1-2i} n_0(v; 1 - i)$. We denote the $S_{1-i} = \{v \in V_n : S_0(v) = 1 - i\}$. By assumption, we have that $|S_{1-i}| < ne^{-\gamma d}$ for some $\gamma \in [0, 1)$.

We will condition on the event of Lemma 2.2.13, which we refer to as $\mathcal{D}_n$. That is, $\mathcal{D}_n$ denotes the event that no more than $\ell_\lambda + 1$ vertices in $L_n(C)$ have a common neighbour. According to Lemma 2.2.13, we have $\mathbb{P}(\mathcal{D}_n) = 1 - o(1)$.

For a partition (U_i, U_{1-i}) of $H_n(C)$, we assume that the vertices in U_j are assigned strategy j, for $j \in \{0, 1\}$. As we pointed out previously, the event $\mathcal{D}_n$ will allow us to ignore the influence of the vertices in $L_n(C)$ on the evolution of those in $H_n(C)$, provided that C is sufficiently large. To this end, we will say that a vertex is i-atypical *with respect to* (U_i, U_{1-i}) if $n_0(v; i) \le \lambda^{1-2i} n_0(v; 1-i)$ for any initial assignment of strategies to the vertices of $L_n(C)$. We will show that a.a.s. for all configurations (U_i, U_{1-i}) of H_n with $|U_{1-i}| < ne^{-\gamma d}$ there is no collection of $\varepsilon|U_{1-i}|$ vertices in $H_n(C)$ which are i-atypical with respect to (U_i, U_{1-i}). This will imply that $G(n, p)$ is such that for any initial configuration (U_i, U_{1-i}) on $H_n(C)$, with $|U_{1-i}| < ne^{-\gamma d}$, and an arbitrary configuration for the vertices in $L_n(C)$, there will be at most $\varepsilon|U_{1-i}|$ vertices in $H_n(C)$ that will adopt strategy $1-i$ in the subsequent round.

We wish to bound the number of vertices in $H_n(C)$ which are i-atypical with respect to a given configuration (U_i, U_{1-i}); thus we define $\hat{S}_i = \{v \in U_i : n_0(v; 1-i) \ge \lambda^{2i-1} n_0(v; i)\}$ and $\hat{S}_{1-i} = \{v \in U_{1-i} : n_0(v; 1-i) \ge \lambda^{2i-1} n_0(v; i)\}$. If there are at least $\varepsilon|S_{1-i}|$ vertices which are i-atypical with respect to (U_i, U_{1-i}), then either $|\hat{S}_i| \ge \varepsilon|U_{1-i}|/2$ or $|\hat{S}_{1-i}| \ge \varepsilon|U_{1-i}|/2$. We will show that

$$\mathbb{P}\left(\bigcup_{1 \le k < ne^{-\gamma d}} \bigcup_{(U_i, U_{1-i}) : |U_{1-i}| = k} \left\{ |\hat{S}_i| \ge \varepsilon|U_{1-i}|/2 \right\} \bigcap \mathcal{D}_n \right) = o(1), \tag{2.18}$$

$$\mathbb{P}\left(\bigcup_{1 \le k < ne^{-\gamma d}} \bigcup_{(U_i, U_{1-i}) : |U_{1-i}| = k} \left\{ |\hat{S}_{1-i}| \ge \varepsilon|U_{1-i}|/2 \right\} \bigcap \mathcal{D}_n \right) = o(1). \tag{2.19}$$

We will show that the union bound indeed suffices to show (2.18) and (2.19). So, firstly, we will consider a fixed partition (U_i, U_{1-i}) as above. To bound $\mathbb{P}(|\hat{S}_i| \ge \varepsilon|U_i|/2)$ and $\mathbb{P}(|\hat{S}_{1-i}| \ge \varepsilon|U_i|/2)$, we translate the defining conditions of $\hat{S}_i$ and $\hat{S}_{1-i}$ into a condition on the degree of these vertices in U_{1-i}. On the event $\mathcal{D}_n$, there are at most $\ell_\lambda + 1$ neighbours of v in L_n. Consider the degree of v inside U_{1-i}, which we denote by $d_{U_{1-i}}(v)$. Similarly, we denote by $d_{S_{1-i} \cap L_n}(v)$ its degree inside $S_{1-i} \cap L_n$. Thus, $d_{U_{1-i}}(v) + d_{S_{1-i} \cap L_n}(v) = n_0(v; 1-i)$.

But since $\mathcal{D}_n$ is realised, we have $d_{S_{1-i}\cap L_n}(v) \leq \ell_\lambda + 1$, whereby

$$d_{U_{1-i}}(v) + \ell_\lambda + 1 \geq n_0(v; 1-i). \tag{2.20}$$

Furthermore, $n_0(v;i) + n_0(v;1-i) = d(v) \geq C$. If $\lambda^{2i-1} n_0(v;i) \leq n_0(v;1-i)$, then

$$\frac{\lambda^{2i-1}+1}{\lambda^{2i-1}} n_0(v;1-i) \geq C.$$

We further bound $n_0(v;1-i)$ using (2.20) and get

$$\frac{\lambda^{2i-1}+1}{\lambda^{2i-1}} \left(d_{U_{1-i}}(v) + \ell_\lambda + 1\right) \geq C.$$

Rearranging this we deduce that

$$d_{U_{1-i}}(v) \geq \frac{\lambda^{2i-1}}{\lambda^{2i-1}+1} C - (\ell_\lambda + 1) \geq \frac{\lambda^{2i-1}}{2(\lambda^{2i-1}+1)} C =: \psi_\lambda C,$$

provided that C is large enough. To summarise, we have proved that if $\lambda^{2i-1} n_0(v;i) \leq n_0(v;1-i)$ and $\mathcal{D}_n$ is realised, then

$$d_{U_{1-i}}(v) \geq \psi_\lambda C. \tag{2.21}$$

We will start with (2.18). Using (2.21), we see that the event in (2.18) is included in the following event: there are disjoint set sets S, U with $1 \leq |U| < ne^{-\gamma d}$ and $|S| = \varepsilon |U|/2$ such that for any $v \in S$ we have $d_U(s) \geq \psi_\lambda C$. Let us consider a set U with $1 \leq |U| \leq ne^{-\gamma d}$ and let $S \subset V_n \setminus U$ be such that $|S| = \varepsilon |U|/2$.

$$\mathbb{P}(\forall v \in S,\ d_U(v) \geq \psi_\lambda C) = \prod_{v \in S} \mathbb{P}(d_U(v) \geq \psi_\lambda C), \tag{2.22}$$

since these are events depending on pairwise disjoint sets of edges. We now observe that for any $v \in S$ the random variable $d_U(v)$ is stochastically dominated by a random variable with distribution $\mathrm{Bin}(|U|, d/n)$. Using that $\binom{n}{k} \le (en/k)^k$, we can bound the above probability in the following way:

$$\mathbb{P}(d_U(v) \ge \psi_\lambda C) \le \binom{|U|}{\psi_\lambda C}\left(\frac{d}{n}\right)^{\psi_\lambda C} \le \left(\frac{ed|U|}{\psi_\lambda n}\right)^{\psi_\lambda C}.$$

Substituting this bound into (2.22), we finally get

$$\begin{aligned}\mathbb{P}(\forall v \in S,\ d_U(v) \ge \psi_\lambda C) &\le \left(\frac{ed|U|}{\psi_\lambda n}\right)^{\psi_\lambda |S| C} \\ &\overset{d \le \alpha(\lambda)\log n}{=} \exp\left(-\psi_\lambda |S| C \log(n/|U|)(1+o(1))\right).\end{aligned}$$

Now we can bound

$$\begin{aligned}&\mathbb{P}\left(\exists S : |S| = \varepsilon|U|/2,\ \forall v \in S,\ d_{U_{1-i}}(v) \ge \psi_\lambda C\right) \\ &\le \binom{n}{\varepsilon|U|/2} \cdot \exp\left(-\psi_\lambda \frac{\varepsilon}{2}|U| C \log\left(\frac{n}{|U|}\right)(1+o(1))\right) \\ &\le \left(\frac{ne}{\varepsilon|U|/2}\right)^{\varepsilon|U|/2} \cdot \exp\left(-\psi_\lambda \frac{\varepsilon}{2}|U| C \log\left(\frac{n}{|U|}\right)(1+o(1))\right) \\ &\le \exp\left(-(\psi_\lambda C - 1)\psi_\lambda \frac{\varepsilon}{2}|U| \log\left(\frac{n}{|U|}\right)(1+o(1))\right).\end{aligned}$$

We are now ready to show (2.18). We write

$$\begin{aligned}&\mathbb{P}\left(\exists U, S : S \cap U = \varnothing,\ 1 \le |U| < ne^{-\gamma d},\ |S| = \varepsilon|U|/2,\ \forall v \in S,\ d_{U_{1-i}}(v) \ge \psi_\lambda C\right) \le \\ &\sum_{1 \le k < ne^{-\gamma d}} \binom{n}{k} \exp\left(-(\psi_\lambda C - 1)\frac{\varepsilon}{2} k \log(n/k)(1+o(1))\right) \\ &\le \sum_{1 \le k < ne^{-\gamma d}} \left(\frac{ne}{k}\right)^k \exp\left(-(\psi_\lambda C - 1)\frac{\varepsilon}{2} k \log(n/k)(1+o(1))\right) \\ &= \sum_{1 \le k < ne^{-\gamma d}} \exp\left(k \log(n/k)(1 - (\psi_\lambda C - 1)\varepsilon/2)(1+o(1))\right) \\ &\overset{\log(n/k) > \gamma d}{\le} \sum_{1 \le k < ne^{-\gamma d}} \exp\left(\gamma d k (1 - (\psi_\lambda C - 1)\varepsilon/2)(1+o(1))\right) = o(1),\end{aligned} \tag{2.23}$$

provided that C is large enough, depending on ε, λ.

Now, we turn to (2.19). Consider a partition (U_i, U_{1-i}) of $H_n(C)$ with $|U_{1-i}|$ as specified above. If $|\hat{S}_{1-i}| > \varepsilon|U_{1-i}|/2$ and $\mathcal{D}_n$ is realised, then there exists a set of $\varepsilon|U_{1-i}|/2$ vertices in U_{1-i}, whose degree inside U_{1-i} is at least $\psi_\lambda C$. Hence, the total degree of this set inside U_{1-i} must be at least $\frac{\varepsilon\psi_\lambda}{2}|U_{1-i}|C$. In turn, the number of edges in U_{1-i} is at least $\frac{\varepsilon\psi_\lambda}{4}|U_{1-i}|C$. Thus, if $e(U_{1-i})$ denotes the number of edges inside S_{1-i} we have

$$\mathbb{P}(|\hat{S}_{1-i}| \geq \varepsilon|U_{1-i}/2|, \mathcal{D}_n) \leq \mathbb{P}\left(e(U_{1-i}) \geq \frac{\varepsilon\psi_\lambda}{4}|U_{1-i}|C\right).$$

We will show that a.a.s. any subset $U \subset V_n$ with $1 \leq |U| < ne^{-\gamma d}$ has $e(U) \leq \frac{\varepsilon\psi_\lambda}{4}|U|C$. Now, $e(U)$ is stochastically dominated from above by a binomially distributed random variable $Y \sim \mathrm{Bin}(|U|^2, d/n)$. So

$$\mathbb{P}\left(e(U) \geq \frac{\varepsilon\psi_\lambda}{4}|U|C\right) \leq \mathbb{P}\left(Y \geq \frac{\varepsilon\psi_\lambda}{4}|U|C\right).$$

We now bound the last probability as follows:

$$\begin{aligned}
\mathbb{P}\left(Y \geq \frac{\varepsilon\psi_\lambda}{4}|U|C\right) &\leq \binom{|U|^2}{\frac{\varepsilon\psi_\lambda}{4}|U|C} \cdot \left(\frac{d}{n}\right)^{\frac{\varepsilon\psi_\lambda}{4}|U|C} \\
&\leq \left(\frac{|U|^2 e}{\frac{\varepsilon\psi_\lambda}{4}|U|C} \cdot \frac{d}{n}\right)^{\frac{\varepsilon\psi_\lambda}{4}|U|C} \\
&\leq \left(\frac{4e}{\varepsilon\psi_\lambda} \cdot \frac{d}{C} \cdot \frac{|U|}{n}\right)^{\frac{\varepsilon\psi_\lambda}{4}|U|C}.
\end{aligned}$$

Since $d = O(\log n)$, we conclude that

$$\mathbb{P}\left(e(U) \geq \frac{\varepsilon\psi_\lambda}{4}|U|C\right) \leq \exp\left(-(1+o(1))\frac{\varepsilon\psi_\lambda}{4}|U|C\log(n/|U|)\right).$$

Arguing as in the case of (2.18), we take the union bound over all choices of the subset U which satisfy the assumed conditions and a similar calculation as in (2.23) (the only difference being that $\varepsilon/2$ is replaced by $\varepsilon/4$) yields:

$$\mathbb{P}\left(\bigcup_{1\le k<ne^{-\gamma d}} \bigcup_{(U_i,U_{1-i}):|U_{1-i}|=k} \left\{|\hat{S}_{1-i}| \ge \varepsilon|U_{1-i}|/2\right\} \bigcap \mathcal{D}_n\right) \le$$
$$\mathbb{P}\left(\exists U : 1 \le |U| < ne^{-\gamma d},\ e(U) \ge \frac{\varepsilon\psi_\lambda}{4}|U|C\right) = o(1).$$

□

We now proceed with the proofs of Theorems 1.1.2, 1.1.3 and 1.1.4.

Proof of Theorem 1.1.2. Let us first point out that by Lemma 2.2.10, if $d \ge c\log n$ with $c > \alpha(\lambda)$, where $\alpha(\lambda)$ is as in the statement of that lemma, then a.a.s. $\mu_1 = 0$; so the last part of Theorem 1.1.2 follows. Hence, we now assume that $d \le \alpha(\lambda)\log n$. We say that $G(n, d/n)$ has the *minority decline property* for some $\gamma > 0$, if whenever the node system $\mathcal{I} = (G(n, d/n), Q, \mathcal{S})$ is such that $\mu_0 \le ne^{-\gamma d}$, then $|m_1 \cap H_n(C)| \le |m_0 \cap H_n(C)|/10$. Observe now that if $G(n, d/n)$ has the minority decline property, and $\mathcal{I} = (G(n, p, Q, \mathcal{S})$ is a node system with $\mu_0 < ne^{-\gamma d}$, for some $\gamma > 0$, then the vertices in H_n will reach unanimity in a finite number of rounds, by repeated applications of this definition.

By Lemma 2.2.10, a.a.s. $\mu_1 < ne^{-\gamma d}$ for some $\gamma > 0$. But by Lemma 2.2.16, if $C > C_{1/10,\lambda}$, then a.a.s. $G(n, p)$ has the minority decline property for $\gamma > 0$ as above. Thus, for every $t \ge 1$ we have $|m_t \cap H_n(C)| \le |m_0 \cap H_n(C)|10^{-t}$. Hence, for $R = \lceil (1/\log 10)\log(|m_1 \cap H_n(C)|)\rceil + 1 = O(\log n)$ we have

$$|m_R \cap H_n| \le |m_1 \cap H_n(C)|10^{-R} < 1.$$

So $|m_R \cap H_n(C)| = 0$. We now show that if C is sufficiently large and every vertex in $H_n(C)$ does not have too many neighbours inside $L_n(C)$, then once $H_n(C)$ has reached unanimity, it will stay there. In particular, Lemma 2.2.13 states that a.a.s. no $\ell_\lambda + 2$ vertices in $L_n(C)$ have a common neighbour. Let us denote this event by $\mathcal{D}_n$. Thus, on $\mathcal{D}_n$ all vertices in $H_n(C)$ have at most $\ell_\lambda + 1$ neighbours inside $L_n(C)$.

Claim 2.2.17. *If $C \geq \max\{\ell_\lambda + 4, (\ell_\lambda + 1)^2\}$ and the vertices in $H_n(C)$ are unanimous just after step t, playing i^* when in the majority regime, then on the event $\mathcal{D}_n$, the vertices of $H_n(C)$ will stay unanimous after step $t+1$.*

Proof of Claim 2.2.17. Indeed, suppose that at step t all vertices of $H_n(C)$, for C to be determined, are unanimous at playing strategy $i \in \{0,1\}$. Consider a vertex $v \in H_n(C)$. If the event $\mathcal{D}_n$ is realised, then all but at most $\ell_\lambda + 1 = \lceil \max\{\lambda, \lambda^{-1}\} \rceil + 1$ neighbours of v play strategy i. So $n_t(v;i) \geq C - (\ell_\lambda + 1)$ and $n_t(v; 1-i) \leq \lceil \max\{\lambda, \lambda^{-1}\} \rceil + 1 < \max\{\lambda, \lambda^{-1}\} + 2$. Suppose first that we are in the majority regime. Then in this case $i = i^*$, by assumption. Vertex v will change strategy if $n_t(v; i^*) < \lambda^{1-2i^*} n_t(v; 1-i^*)$ (cf. (2.4)). But

$$\lambda^{1-2i^*} n_t(v; 1-i^*) < \lambda^{1-2i^*}(\ell_\lambda + 1) \leq \lambda^{1-2i^*}(\max\{\lambda, \lambda^{-1}\} + 2) \overset{(2.15)}{<} 1 + 2 = 3.$$

Therefore must have that $C - (\ell_\lambda + 1) < 3$. However, choosing $C \geq \ell_\lambda + 4$ leads to a contradiction.

If we are in the minority regime, then we want to show that v will switch strategy. By (2.5) if the entire set $H_n(C)$ plays strategy i^*, then $v \in H_n(C)$ will switch strategy if $n_t(v; i^*) > \lambda^{1-2i^*} n_t(v; 1-i^*)$. But $n_t(v; i^*) \geq C - (\ell_\lambda + 1)$ and $n_t(v; 1-i) \leq \ell_\lambda + 1$. As seen above, if $C \geq \ell_\lambda + 4$ then $n_t(v; i^*) \geq C - (\ell_\lambda + 1) \geq 3 > \lambda^{1-2i^*}(\ell_\lambda + 1) \geq n_t(v; 1-i^*)$. Now, if the entire set $H_n(C)$ plays strategy $1-i^*$, then $v \in H_n(C)$ will switch strategy if $n_t(v; 1-i^*) > \lambda^{1-2(1-i^*)} n_t(v; i^*)$. But $n_t(v; i^*) \leq \ell_\lambda + 1$ whereas $n_t(v; 1-i^*) \geq C - (\ell_\lambda + 1)$. Thus v will switch strategy if $C - (\ell_\lambda + 1) \geq \lambda^{2i^*-1}(\ell_\lambda + 1)$. Therefore, choosing $C \geq (\ell_\lambda + 1)^2$ yields a contradiction and completes the proof of the claim. □

Remark 2.2.18. *For the minority regime, the above claim and Lemma 2.2.10 imply that, when unanimity occurs within $H_n(C)$, its vertices will be playing strategy $1-i^*$ at odd steps and strategy i^* at even steps. In the majority regime, they stabilise to strategy i^*.*

Note that by Claim 2.2.11, if $d \gg 1$, then for any fixed $C \in \mathbb{N}$ we have a.a.s. $|H_n(C)| \geq n(1 - o(1))$. Therefore, the above analysis implies that for any $\varepsilon > 0$ there exists $\beta =$

$\beta(\varepsilon, \lambda) > 0$ such that if $d \gg 1$, then a.a.s. at least $n(1-\varepsilon)$ vertices in $G(n, d/n)$ will be unanimous after at most $\beta \log n$ rounds. $\square$

Proof of Theorem 1.1.3. Suppose first that $d = c_\lambda \log n + \log\log n + \omega(n)$, where $\omega(n) \to -\infty$ as $n \to +\infty$. By Lemma 2.2.8 *i.*, a.a.s. there are $(\ell_\lambda, 1)$-blocking stars in $L_1(G(n, d/n))$ that are set to the $(1-i^*, 1-i^*)$-configuration. By Claim 2.2.3, those will retain this configuration forever and, therefore, they will be in disagreement with the vertices in $H_n(C)$. Hence, $u_n^{(1)}$, the probability that $L_1(G(n, d/n))$ becomes eventually unanimous, tends to 0 as $n \to +\infty$. We will now consider the cases where $d = c_\lambda \log n + \log\log n + \omega(n)$, where either $\omega(n) \to +\infty$ or $\omega(n) \to c \in \mathbb{R}$, as $n \to +\infty$. We will need the following claim.

Claim 2.2.19. *Let Q be in the majority regime. If a vertex v has at most $\ell_\lambda - 1$ neighbours playing strategy $1 - i^*$ but at least one playing strategy i^*, it will play strategy i^* in the next round.*

Proof of Claim 2.2.19. If v already plays strategy i^*, then it will change strategy if $n_t(v; i^*) < \lambda^{1-2i^*} n_t(v; 1-i^*)$. But $n_t(v; i^*) \geq 1$ and $n_t(v; 1-i^*) \leq \ell_\lambda - 1$. Hence we have,

$$\lambda^{1-2i^*} n_t(v; 1-i^*) \leq \lambda^{1-2i^*}(\ell_\lambda - 1) < \lambda^{1-2i^*} \cdot \max\{\lambda, \lambda^{-1}\} \overset{(2.15)}{=} 1.$$

So v will not change strategy. Now, if v already plays strategy $1-i^*$, then it will not change its strategy if $n_t(v; 1-i^*) \geq \lambda^{1-2(1-i^*)} n_t(v; i^*) = \lambda^{2i^*-1} n_t(v; i^*)$. But $n_t(v; 1-i^*) \leq \ell_\lambda - 1$ whereas $\lambda^{2i^*-1} n_t(v; i^*) \geq \lambda^{2i^*-1} > \ell_\lambda - 1$. Therefore, v will change its strategy into i^*. $\square$

Suppose that $d = c_\lambda \log n + \log\log n + \omega(n)$ with $\omega(n) \to +\infty$ as $n \to +\infty$. Let $G^{(L)}(n, d/n)$ denote the subgraph of $G(n, d/n)$ induced by the vertices in $L_n(C)$. By Lemma 2.2.14 (with $\ell = \ell_\lambda$) and Lemma 2.2.15, every connected component of $G^{(L)}(n, d/n)$ is a tree of order at most $\ell_\lambda + 1$. Let T be one of these connected components that is a subgraph of $L_1(G(n, d/n))$.

If $|T| \leq \ell_\lambda$, then all its vertices have degree at most $\ell_\lambda - 1$ in T. For $i \geq 1$, let $T^{(i)}$ denote the set of vertices in T that are at distance i from $H_n(C)$. Once the vertices in

$H_n(C)$ have been unanimous on strategy i^*, they will stay there forever. By Claim 2.2.19 the vertices in $T^{(1)}$ will adopt strategy i^* eventually and remain there forever. Assuming that the vertices $T^{(i)}$ have adopted strategy i^* forever, then the vertices of $T^{(i+1)}$ will adopt strategy i^* too (provided it is non-empty) by Claim 2.2.19 and remain there forever. Hence, the entire vertex set of T will adopt strategy i^*.

Suppose now that $|T| = \ell_\lambda + 1$. If all its vertices have degree at most $\ell_\lambda - 1$, then eventually the vertices of T adopt strategy i^*, by the above argument. If there is a vertex in T of degree ℓ_λ within T, then T must be a star. However, by Lemma 2.2.8 *ii.* a.a.s. this is not an $(\ell_\lambda, 1)$-blocking star. Thus, one of its leaves must have a neighbour in $H_n(C)$. Since it has degree 1 ($\leq \ell_\lambda - 1$) inside T, then by Claim 2.2.19 it adopts strategy i^* after $H_n(C)$ becomes unanimous and stays there forever. Subsequently, the centre of T will do so (it also has at most $\ell_\lambda - 1$ neighbours that are not playing strategy i^*) and finally, the remaining leaves adopt it as well. Moreover, the above argument shows that the only connected components of $G^{(L)}(n, d/n)$ which may not adopt strategy i^* are the $(\ell, 1)$-blocking stars, for $\ell \leq \ell_\lambda$. If there are no $(\ell_\lambda, 1)$-blocking stars, then $L_1(G(n, d/n))$ will then become unanimous. Firstly, let us observe that if $\omega(n) \to c \in \mathbb{R}$ as $n \to \infty$, then by Lemma 2.2.5 *iii.* a.a.s. there are no $(\ell_\lambda + 1, 1)$-blocking stars as

$$\mathbb{P}(X_{\ell_\lambda+1,1,n} > 0) \leq 2e^{-(\ell_\lambda+1)\Omega(\log n)} = o(1).$$

Furthermore, by Lemma 2.2.13 a.a.s. there are no $\ell_\lambda + 2$ vertices of degree 1 that have a common neighbour. Therefore, a.a.s. there are no $(\ell, 1)$-blocking stars for any $\ell \geq \ell_\lambda + 1$. Now, by Lemma 2.2.8 *ii.*, the random variable $X^{(1)}_{\ell_\lambda,1,n}$ converges in distribution as $n \to +\infty$ to a random variable distributed as $\mathrm{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!)$. Thus, for any integer $k \geq 0$, we have

$$\mathbb{P}\left(X^{(1)}_{\ell_\lambda,1,n} = k\right) \to \mathbb{P}\left(\mathrm{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!) = k\right),$$

as $n \to +\infty$.

Suppose now that $X^{(1)}_{\ell_\lambda,1,n} = k$, for some $k \in \mathbb{N}_0$. The case $k = 0$ was treated above and unanimity is attained a.a.s. (on the conditional space where $X^{(1)}_{\ell_\lambda,1,n} = 0$). Let us consider the case $k \geq 1$. If an $(\ell_\lambda, 1)$-blocking star is initially set to $(1 - i^*, 1 - i^*)$-configuration, then by Claim 2.2.3 it will stay in this configuration forever. We thus conclude that if unanimity is achieved then no $(\ell_\lambda, 1)$-blocking star is initially set to $(1 - i^*, 1 - i^*)$-configuration. The probability of this is $(1 - 1/2^{\ell_\lambda+1})^k$. Also, if all $(\ell_\lambda, 1)$-blocking stars attached to $L_1(G(n, d/n))$ are initially set to (i^*, i^*)-configuration, they will remain so forever (cf. Claim 2.2.4) and will be synchronised with the vertices of $H_n(C)$. Thus, $L_1(G(n, d/n))$ will be unanimous. The probability of this is $1/2^{k(\ell_\lambda+1)}$.

Consequently,

$$\limsup_{n\to+\infty} u_n^{(1)} \leq \sum_{k=0}^{\infty} \left(1 - \frac{1}{2^{\ell_\lambda+1}}\right)^k \mathbb{P}\left(\text{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!) = k\right).$$

and

$$\liminf_{n\to+\infty} u_n^{(1)} \geq \sum_{k=0}^{\infty} \left(\frac{1}{2^{\ell_\lambda+1}}\right)^k \mathbb{P}\left(\text{Po}(e^{c(\ell_\lambda+1)}/\ell_\lambda!) = k\right).$$

Since $H_n(C)$ will reach unanimity in at most $\beta \log n$ steps, the above case analysis implies that $L_1(G(n, d/n))$ will reach unanimity in at most $\beta \log n + O(1)$ steps. □

Proof of Theorem 1.1.4. Let us recall that $\ell'_\lambda = \lfloor \max\{\lambda, \lambda^{-1}\} \rfloor = \lfloor \lambda^{2i^*-1} \rfloor$. Suppose first that $d = \frac{1}{2}\log n + \frac{1+\ell'_\lambda}{2} \log\log n + \omega(n)$, where $\omega(n) \to -\infty$ as $n \to +\infty$. By Lemma 2.2.8 *i.* (setting $\ell = 1$ and $k = \ell'_\lambda$ therein), a.a.s. there are $(1, \ell'_\lambda)$-blocking stars in $L_1(G(n, d/n))$ that are initially set to the $(i^*, 1 - i^*)$-configuration. By Claim 2.2.1, those will retain this configuration forever and, therefore, they will be in disagreement with the vertices in $H_n(C)$. Hence, $u_n^{(1)}$, the probability that $L_1(G(n, d/n))$ becomes eventually unanimous, tends to 0 as $n \to +\infty$.

Now, suppose that $d = \frac{1}{2}\log n + \frac{1+\ell'_\lambda}{2} \log\log n + \omega(n)$, where $\omega(n) \to +\infty$. As before, we let $G^{(L)}(n, d/n)$ denote the subgraph of $G(n, d/n)$ induced by the vertices in $L_n(C)$. By Lemma 2.2.14 (with $\ell = 1$) a.a.s. every connected component of $G^{(L)}(n, d/n)$ is of order at most 2. That is, a.a.s. every component of $G^{(L)}(n, d/n)$ is either a vertex or an edge.

Let T be one of these connected components that is a subgraph of $L_1(G(n, d/n))$. If T is a vertex, then it will synchronise with the vertices $H_n(C)$ after the R^{th} step, where the vertices of $H_n(C)$ arrive at unanimity. Thus, all its neighbours (which lie in $H_n(C)$) will play the same strategy, say i, by (2.5) this vertex will adopt strategy $1-i$ in the next round and be in agreement with the vertices of $H_n(C)$ (cf. Claim 2.2.17).

Suppose now that T is an edge with one of its endpoints being adjacent to vertices in $H_n(C)$. Hence, T is a $(1, k)$-blocking star for some $k \in \mathbb{N}$. But in fact, $k > \ell'_\lambda$ as $\omega(n) \to +\infty$ and by Lemma 2.2.8 *iii*. a.a.s. there are no $(1, k)$-blocking stars with $k \le \ell'_\lambda$. Such a $(1, k)$-blocking star with $k > \ell'_\lambda$, will have its k connectors inside $H_n(C)$. But recall that these will arrive at unanimity after step R and will start alternating simultaneously between states i^* and $1 - i^*$. So by Claim 2.2.2, the $(1, k)$-blocking star will synchronise with them.

Finally, suppose that $T = v_1v_2$ is an edge where both its endpoints v_1 and v_2 have at least one neighbour in $H_n(C)$. Let $t \ge R$ be a step at which $S_t(v) = i^*$, for all $v \in H_n(C)$. Assume that v_1 and v_2 are not unanimous with $H_n(C)$. In particular, suppose that $S_t(v_1) = S_t(v_2) = 1-i^*$. Vertex v_1 will not switch strategy, if $n_t(v_1; 1-i^*) \le \lambda^{2i^*-1}n_t(v_1; i^*)$. But $n_t(v_1; 1 - i^*) = 1$, $n_t(v_1; i^*) \ge 1$ and $\lambda^{2i^*-1} > 1$. So the inequality is satisfied. The same holds for v_2. Thereby, $S_{t+1}(v_1) = S_{t+1}(v_2) = 1 - i^*$ and as $S_{t+1}(v) = 1 - i^*$ for all $v \in H_n(C)$, thereafter v_1, v_2 will be synchronised with $H_n(C)$.

Suppose now that $S_t(v_1) = 1 - i^*$ but $S_t(v_2) = i^*$. Then $S_{t+1}(v_1) = 1 - i^*$ as v_1 has no neighbours who play strategy $1 - i^*$. Also, $S_{t+1}(v_2) = 1 - i^*$, since $n_t(v_2; i^*) > \lambda^{1-2i^*}n_t(v_2; 1-i^*)$. The latter holds since $n_t(v_2; i^*) \ge 1$, $n_t(v_2; 1-i^*) = 1$ and $\lambda^{1-2i^*} < 1$. As $S_{t+1}(v) = 1 - i^*$ for all $v \in H_n(C)$, thereafter v_1, v_2 will stay synchronised with $H_n(C)$. By symmetry, an analogous argument can be used for the case $S_t(v_1) = i^*$ but $S_t(v_2) = 1 - i^*$.

We thus conclude that if $\omega(n) \to \infty$, then $u_n^{(1)} \to 1$ as $n \to \infty$.

Finally, suppose that $\omega(n) \to c \in \mathbb{R}$ as $n \to \infty$. By Lemma 2.2.8 *ii*., the random variable $X^{(1)}_{1,\ell'_\lambda,n}$ converges in distribution as $n \to +\infty$ to a random variable distributed as

$\text{Po}(e^{2c}/\ell'_\lambda!)$. Thus, for any integer $k \geq 0$, we have

$$\mathbb{P}\left(X^{(1)}_{1,\ell'_\lambda,n} = k\right) \to \mathbb{P}\left(\text{Po}(e^{2c}/\ell'_\lambda!) = k\right)$$

as $n \to +\infty$.

Suppose now that $X^{(1)}_{1,\ell'_\lambda,n} = k$, for some $k \in \mathbb{N}_0$. The case $k = 0$ was treated above and unanimity is attained a.a.s. (on the conditional space where $X^{(1)}_{1,\ell'_\lambda,n} = 0$). Indeed as no $(1, \ell'_\lambda)$-blocking stars are present, and hence we have that,

$$\liminf_{n\to+\infty} u^{(1)}_n \geq \mathbb{P}\left(\text{Po}(e^{2c}/\ell'_\lambda!) = 0\right).$$

Let us consider the case $k \geq 1$. If an $(1, \ell'_\lambda)$-blocking star is initially set to $(i^*, 1 - i^*)$-configuration, then by Claim 2.2.1 it will stay in this configuration forever. In other words, if unanimity is achieved, then no $(1, \ell'_\lambda)$-blocking star is initially set to $(i^*, 1 - i^*)$-configuration. The probability of this is $(1 - 1/4)^k = (3/4)^k$. Consequently,

$$\limsup_{n\to+\infty} u^{(1)}_n \leq \sum_{k=0}^{\infty} \left(\frac{3}{4}\right)^k \mathbb{P}\left(\text{Po}(e^{2c}/\ell'_\lambda!) = k\right).$$

Since $H_n(C)$ will reach unanimity in at most $\beta \log n$ steps, the above argument implies that $L_1(G(n, d/n))$ will reach unanimity in at most $\beta \log n + O(1)$ steps. □

We conclude this section with the proof of Lemma 2.2.16.

2.3 Unbiased Node Systems in Dense Regimes

In this chapter, we explore the remaining cases that are not covered in Theorem 1.1.1. We consider the range for $p \geq \Lambda n^{-1/2}$, for large positive Λ and payoff skew $\lambda = 1$. We have two cases depending on the regime of Q. We note that if Q is in the majority regime, then this is covered by the main result of [32]. Therefore the majority of this section focuses on when Q is in the minority regime. By considering a coupling of the minority regime to a suitably designed majority game, we can confirm that an analogous result for the minority regime result holds.

2.3.1 Majority and Minority Dynamics

We recall that $S_t(v)$ is the state of a vertex v at a discrete time-step $t \geq 0$. We consider a process running on a suitably dense realisation of $G(n,p)$, and also utilise the initial configuration $\mathcal{S}_{1/2}$. If our interacting node system $(G(n,p), Q, \mathcal{S}_{1/2})$ is in the majority regime with $\lambda(Q) = 1$, then its evolution coincides with the *majority dynamics process*. The latter is defined by the following evolution rule:

$$S_{t+1}(v) = \begin{cases} 1 & \text{if } n_t(v;1) > n_t(v;0); \\ 0 & \text{if } n_t(v;1) < n_t(v;0); \\ S_t(v) & \text{if } n_t(v;1) = n_t(v;0). \end{cases} \tag{2.24}$$

In other words, in the majority dynamics process, a node will always choose to adopt the state shared by the majority of its neighbours. If there is a tie, then its state will remain unchanged. Goles and Olivos [39] showed that majority dynamics on a finite graph becomes eventually periodic with a period at most 2. More specifically, there is a t_0 depending on the graph such that for any $t > t_0$ and for any vertex v we have $S_t(v) = S_{t+2}(v)$. Majority dynamics is also a special case of voting with at least two alternatives; see [64]. Results on the evolution of majority dynamics on the random graph $G(n,p)$ were obtained recently by Benjamini et al. [9]. In [32], Fountoulakis et al. proved the following theorem confirming the rapid stabilisation of the majority dynamics process on a suitably dense $G(n,p)$, confirming a conjecture stated in [9]. Let M_0 be the most popular vertex state seen across the initial configuration.

Theorem 2.3.1 [32]. *For all $\varepsilon \in [0,1)$ there exist Λ, n_0 such that for all $n > n_0$, if $p \geq \Lambda n^{-\frac{1}{2}}$, then $G(n,p)$ is such that with probability at least $1-\varepsilon$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/2}$, the vertices in V_n following the majority dynamics rule, unanimously* *have state M_0 after four rounds.*

We note that Theorem 2.3.1 allows us to conclude the stabilisation of the interacting node system with $\lambda = 1$ in the majority regime. On the other hand, if we consider an interacting node system in the minority regime with $\lambda = 1$, then the process coincides with the *minority dynamics process*, described by the following:

$$S_{t+1}(v) = \begin{cases} 1 & \text{if } n_t(v;1) < n_t(v;0); \\ 0 & \text{if } n_t(v;1) > n_t(v;0); \\ S_t(v) & \text{if } n_t(v;1) = n_t(v;0). \end{cases} \tag{2.25}$$

Under these rules, nodes will update to the state shared by the minority of their neighbours. It can be readily checked from (2.4) and (2.5), respectively, that the evolution of the above systems is identical to an interacting node system with $\lambda = 1$. We show that in the minority regime, unanimity is also achieved within at most four rounds too. However, (2.25) implies that vertex strategies will alternate synchronously with period two. Our theorem concerning the evolution of minority dynamics is analogous to Theorem 2.3.1.

Theorem 2.3.2. *For all $\varepsilon \in (0,1]$ there exist Λ, n_0 such that for all $n > n_0$, if $p \geq \Lambda n^{-\frac{1}{2}}$, then $G(n,p)$ is such that with probability at least $1-\varepsilon$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/2}$, the vertices in V_n following the minority dynamics rule will unanimously have the same state after four rounds.*

Throughout this section, we consider $p = d/n$ where $d \geq \Lambda\sqrt{n}$, for a suitably large constant Λ. Thus, in comparison to the previous section, we will now work in a denser realisation of $G(n,p)$. We will comment on sparser regimes in the discussion section of this chapter. We consider an initial configuration of $\mathcal{S}_{1/2}$, and apply the evolution rules from (2.24) in the majority regime, or the evolution rules from (2.25) in the minority regime. We refer to the node system using evolution rule (2.24) as the *majority game* which we denote as $\left(G(n,p), \mathcal{S}_{1/2}\right)^{>}$; while we refer to the system given by evolution rule (2.25) as the *minority game*, denoted $\big(G(n,p), \mathcal{S}_{1/2}\big)^{<}$. We show that in the minority game, unanimity is achieved after at most four rounds with high probability. As noted above, the majority game will give rise to stability, while the minority game produces a system with period two.

The quantity $\eta_t := \Big||P_t| - |N_t|\Big|$ represents the size of the majority of the dominant strategy at time t. Due to the distribution of the $\mathcal{S}_{1/2}$, we have with probability $1-\varepsilon$ that the quantity $\eta_0 = \Big||P_0| - |N_0|\Big|$ will be sufficiently bounded away from zero.

Lemma 2.3.3 [32]**.** *Given $\varepsilon > 0$, set $c = c(\varepsilon) = \sqrt{2\pi}\varepsilon/20$. Then across the probability space $\mathcal{S}_{1/2}$,*

$$\mathbb{P}\left(\eta_0 \geq 2c\sqrt{n}\right) \geq 1 - \varepsilon/4,$$

for any n sufficiently large.

The proof of this lemma is a direct consequence of the Local Limit Theorem (see Theorem 2.4.3 below). We define $\mathcal{E}_c^+$ to be the event that $|P_0| - |N_0| \geq 2c\sqrt{n}$, and $\mathcal{E}_c^-$ to be the event that $|N_0| - |P_0| \geq 2c\sqrt{n}$. The events $\mathcal{E}_c^+$ and $\mathcal{E}_c^-$ occur with equal probability; therefore by symmetry we may condition on either $\mathcal{E}_c^+$ or $\mathcal{E}_c^-$, without loss of generality. We will consider a selection of results from [32] (Lemmas 3.5, 3.6 therein), which will be useful in the minority regime analysis. The first result concerns the expectation and variance of $n_1(v;1)$, given that $\mathcal{E}_c^+$ has occurred. For a vertex $v \in V_n$ define the event $\mathcal{N}(v) = \{|d(v) - d| < d^{2/3}\}$.

Lemma 2.3.4. *Consider the majority game $M = (G(n,p), \mathcal{S}_{1/2})^{>}$ where $p = d/n$. Let $c = c(\varepsilon)$ be the constant given in Lemma 2.3.3. Then there exists a constant ζ, (independent of ε) such that for any $v \in V_n$ and any n sufficiently large the following holds: for any configuration $s_0 \in \mathcal{E}_c^+$ and any $k \in \mathbb{N}$ such that $|k - d| < d^{2/3}$:*

$$\mathbb{E}\left[n_1(v;1) \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right] \geq \frac{k}{2} + \frac{\zeta c}{7}\left(\frac{d^3}{n}\right)^{1/2}.$$

Moreover, there exists a positive constant α, such that for any $k \in \mathbb{N}$ with $|k - d| < d^{2/3}$

we have

$$\text{Var}\left[n_1(v;1) \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right] \leq \alpha d.$$

By applying Lemma 2.3.4, we now proceed to prove an adjustment to a result from [32]. In the modified result, we show that with high probability a vertex v will have a $n_1(v;1)$ sufficiently bounded away from $d(v)/2$. In [32], the authors show that this quantity is at least $d(v)/2$. However, in order for us to utilise this result for the minority game, we will instead require the stronger bound $n_1(v;1) \geq d(v)/2 + 2\gamma\sqrt{d}$, with high probability, for some positive constant γ. We elaborate on the reasoning behind this assertion in Section 2.4.

Lemma 2.3.5. *Let $\varepsilon > 0$, and c a positive constant as given in Lemma 2.3.3. Consider the majority game $M = (G(n,p), \mathcal{S}_{1/2})^{>}$ where $p = d/n$. For any positive constant γ there exists a positive constant $\Lambda = \Lambda(\gamma, \varepsilon)$, such that for all n sufficiently large if $d \geq \Lambda n^{1/2}$ and $v \in V_n$, the followings holds:*

$$\mathbb{P}\left(n_1(v;1) < \frac{d(v)}{2} + 2\gamma\sqrt{d} \;\middle|\; \mathcal{E}_c^+\right) < \varepsilon.$$

Proof. This argument is a direct application of Chebyshev's inequality. Fix $s_0 \in \mathcal{E}_c^+$ and $k \in \mathbb{N}$ such that $|k - d| < d^{2/3}$. By applying Lemma 2.3.4 and subtracting $2\gamma\sqrt{d}$, from both sides we have that:

$$\left|\mathbb{E}\left[n_1(v;1) \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right] - \left(\frac{d(v)}{2} + 2\gamma\sqrt{d}\right)\right| \geq \left(\frac{d^3}{n}\right)^{1/2} \frac{\zeta c}{7} - 2\gamma\sqrt{d}.$$

By Lemma 2.3.4 we have $\text{Var}\left[n_1(v;1) \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right] \leq \alpha d$. We now apply Chebyshev's inequality to bound the probability that $n_1(v;1) < d(v)/2 + 2\gamma\sqrt{d}$. This gives

$$\mathbb{P}\left(n_1(v;1) < \frac{d(v)}{2} + 2\gamma\sqrt{d} \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right) \leq$$

$$\frac{49\alpha d}{\left(\zeta c\,(d^3/n)^{1/2} - 14\gamma\sqrt{d}\right)^2} \leq \frac{49\alpha}{\zeta c\left(\frac{d}{\sqrt{n}}\right)\left[\zeta c\left(\frac{d}{\sqrt{n}}\right) - 28\gamma\right]}.$$

We now recall that $d > \Lambda\sqrt{n}$. If we take $\Lambda > (1+28\gamma)/(\zeta c)$, then $\zeta c\,(d/\sqrt{n}) - 28\gamma \geq 1$. Applying this inequality to the denominator, and choosing $\Lambda > 2 \cdot \max\{(1+28\gamma)/\zeta c, 49\alpha/\varepsilon c\zeta\}$ we have:

$$\mathbb{P}\left(n_1(v;1) < \frac{d(v)}{2} + 2\gamma\sqrt{d} \mid \mathcal{S}_{1/2} = s_0, d(v) = k\right) \leq \frac{49\alpha}{\Lambda c\zeta} \leq \varepsilon/2.$$

Integrating over all possible choices of $s_0 \in \mathcal{E}_c^+$ and k such that $\mathcal{N}(v)$ is realised, we obtain

$$\mathbb{P}\left(n_1(v;1) < \frac{d(v)}{2} + 2\gamma\sqrt{d} \mid \mathcal{E}_c^+, \mathcal{N}(v)\right) \leq \varepsilon/2.$$

But $\mathbb{P}[\mathcal{N}(v)] = 1 - o(1)$, (which follows by a standard application of the Chernoff bound (1.3)), we deduce that for n sufficiently large we have

$$\mathbb{P}\left(n_1(v;1) < \frac{d(v)}{2} + 2\gamma\sqrt{d} \mid \mathcal{E}_c^+\right) \leq \varepsilon.$$

□

2.4 The Minority Regime

We now work within the minority regime, proving Theorem 2.3.2. Recall that by (2.25), a vertex will update to the state shared across the minority of its neighbours. In the event of a tie, the vertex will remain in its current state.

We observe similarities with Theorem 2.3.1, namely the fact that unanimity occurs, and is achieved within at most four rounds. However, the system is no longer stable but will become periodic with period two after unanimity is reached.

We wish to relate the proof of Theorem 2.3.2 to Theorem 2.3.1. We first start by showing that the first round of the minority game $m = \left(G(n,p), \mathcal{S}_{1/2}\right)^{<}$ can be approximated by a specific majority game which starts on the *complementary configuration*. For a configuration S on V_n we define the *complementary configuration* $\bar{S}$ as follows: for every $v \in V_n$ we set $\bar{S}(v) = 1 - S(v)$. Suppose we have an interacting node system $\mathcal{I} = \left(G, \mathcal{S}_{1/2}\right)^*$, where $* \in \{<,>\}$. For $v \in V_n$, we denote $S_t^{\mathcal{I}}(v)$ to be the strategy of a vertex v in the game $\mathcal{I}$ at time t; similarly we define $n_t^{\mathcal{I}}(v;i) = |\{u : S_t^{\mathcal{I}}(u) = i\} \cap N_G(v)|$. While these definitions are similar to their counterparts, we would like to emphasise the role of $\mathcal{I}$, as we will generally work with two different systems: a minority game $\mathcal{I} = m$ (when $*$ is $<$) and a majority game $\mathcal{I} = M$ (when $*$ is $>$). We now state the following lemma concerning the approximation of a minority game m to a suitably designed majority game M.

Lemma 2.4.1. *Let $G = (V, E)$ be a graph and $\mathcal{S} : V \to \{0, 1\}$ be a configuration. Let $m = (G, \mathcal{S})^<$ be a minority game, and $M = (G, \bar{\mathcal{S}})^>$ be a majority game with initial configuration $\bar{\mathcal{S}}$. If for $v \in V(G)$ we have that $n_0^m(v; 0) \neq n_0^m(v; 1)$, then $S_1^m(v) = S_1^M(v)$. If $n_0^m(v; 0) = n_0^m(v; 1)$, then $S_1^m(v) = 1 - S_1^M(v)$.*

Lemma 2.4.1 allows us to deduce the behaviour of a significant number of vertices in the first round of the minority process. The idea is to take a minority game, complement each of the vertex strategies, and then allow one round of evolution to occur using the majority rules. As long as a vertex satisfies the condition $n_0^m(v; 0) \neq n_0^m(v; 1)$, it will have the same state as if it had just evolved using the minority rules on the original configuration. We refer to the additional condition $n_0^m(v; 0) = n_0^m(v; 1)$, as the *equal neighbourhoods condition* (ENC).

Proof of Lemma 2.4.1. We first assume that $n_0^m(v; 0) \neq n_0^m(v; 1)$. We split our analysis into cases which depend on both the current state of the vertex, along with which of $n_0^m(v; 0)$ and $n_0^m(v; 1)$ is larger. In all four cases the argument is identical; we simply must show that $S_1^m(v) = S_1^M(v)$.

Suppose $S_0^m(v) = 0$, and $n_0^m(v; 0) > n_0^m(v; 1)$. By applying the minority rules, we see that $S_1^m(v) = 1$. We now consider the complementary state, $1 - \bar{S}_0^m(v) = S_0^M(v) = 1$.

As we have that $n_0^m(v;0) > n_0^m(v;1)$, then by the definition of complementary initial configuration, it must be the case that $n_0^M(v;0) < n_0^M(v;1)$. By applying the majority rules to the vertex v we have that $S_1^M(v) = 1$; therefore, $S_1^m(v) = S_1^M(v)$.

For the case where $n_0^m(v;0) = n_0^m(v;1)$, we observe that in both games v has an equal of the number of the vertices playing strategy one and zero in its neighbourhood. Therefore, it follows that $S_0^m(v) = S_1^m(v)$ and $S_0^M(v) = S_1^M(v)$. However, by the definition of complementary states we have that $S_0^m(v) = 1 - S_0^M(v)$, and thus $S_1^m(v) = 1 - S_1^M(v)$. □

The main application of the above lemma is to connect the games m and M on $G(n, d/n)$ with initial configuration $\mathcal{S}_{1/2}$. It is at that point where we apply Lemma 2.3.5 to the game m. However, we must consider which vertices satisfy the equal neighbourhoods condition. For each vertex $v \in V_n$, we say that $v \in \mathrm{ENC}$ if $n_0^m(v;0) = n_0^m(v;1)$, and we define $\mathrm{EQ}(v) = |\{w : w \in \mathrm{ENC}\} \cap N_{G(n,d/n)}(v)|$. Let γ be a positive constant. We say that a vertex $v \in V_n$ has a γ-*decisive neighbourhood* in $G(n, d/n)$ if $\mathrm{EQ}(v) < \gamma\sqrt{d}$, and a vertex has a γ-*abundant neighbourhood* in $G(n, d/n)$ if $n_1^M(v;1) \geq 2\gamma\sqrt{d} + d(v)/2$. We say that a vertex $v \in V_n$ is γ-*good* if v has a γ-abundant and a γ-decisive neighbourhood in $G(n, d/n)$. The following corollary illustrates the role of γ-good vertices.

Corollary 2.4.2. *Let $\gamma > 0$. If a vertex $v \in V_n$ is γ-good in $G(n, d/n)$ for a given initial configuration, then $S_2^m(v) = 0$.*

Proof. Let us abbreviate $G(n, d/n)$ by G. Given an initial configuration on V_n, set up the systems m and M as in Lemma 2.4.1, and assume that $v \in V_n$ is γ-good vertex. We show that $n_1^m(v;1) > d(v)/2$; we proceed by applying Lemma 2.4.1 to bound $n_1^m(v;1)$ from below. This lemma implies that if $u \notin \mathrm{ENC}$, then $S_1^m(u) = S_1^M(u)$. Hence, for all $u \in N_G(v) \setminus \mathrm{ENC}$, if $S_1^M(u) = 1$, then $S_1^m(u) = 1$. However for $w \in \mathrm{ENC} \cap N_G(v)$, we have that $S_1^m(w) = 1 - S_1^M(w)$. If we assume that for all $w \in \mathrm{ENC} \cap N_G(v)$ we have that $S_1^M(w) = 1$, then we would minimise the size of $n_1^m(v;1)$ as in that case $S_1^m(w) = 0$. As a direct consequence, we have that $n_1^m(v;1) \geq n_1^M(v;1) - \mathrm{EQ}(v)$. Therefore, by combining this bound with the definitions of γ-abundance and γ-decisiveness, we have:

$$\begin{aligned} n_1^m(v;1) &\geq n_1^M(v;1) - \mathrm{EQ}(v) > n_1^M(v;1) - \gamma\sqrt{d} \geq \frac{d(v)}{2} + 2\gamma\sqrt{d} - \gamma\sqrt{d} \\ &= \frac{d(v)}{2} + \gamma\sqrt{d}. \end{aligned} \qquad \square$$

Hence we have that v changes to strategy zero after two rounds.

2.4.1 Bounding the Size of $\mathrm{EQ}(v)$

In light of Corollary 2.4.2, our aim is to show that there are a significant number of good vertices. We first show that with high probability, there are a sufficient number of vertices with a γ-decisive neighbourhood. The proof of this bound will require us to invoke the Local Limit Theorem for sums of Bernoulli-distributed random variables, which follows from Theorem V.4 p.111 in [77]. We state this result as follows.

Theorem 2.4.3 [77]**.** *Let $X_1, \dots X_n$ be independent identically distributed Bernoulli-distributed random variables with $\mathbb{E}(X_1) = \mu$ and $\mathrm{Var}(X_1) = \mu - \mu^2 =: \sigma^2 > 0$. Let also $X = \sum_{i=1}^n X_i$. There exists ρ depending on μ for which:*

$$\sup_{i \in \mathbb{N}_0} \left| \sqrt{\mathrm{Var}[X]} \cdot \mathbb{P}[X = i] - \frac{1}{2\pi} \exp\left(-\frac{(i - \mathbb{E}[X])^2}{2\mathrm{Var}[X]} \right) \right| < \frac{\rho}{\sqrt{\mathrm{Var}[X]}}.$$

The following lemma provides an upper bound on the number of vertices in the neighbourhood of $v \in V_n$ in $G(n,p)$ which satisfy the equal neighbourhoods condition.

Lemma 2.4.4. *Consider $G(n,p)$ with $np =: d = d(n) \to \infty$ as $n \to \infty$ and let $v \in V_n$. For every $\varepsilon > 0$, there exist positive constants γ and n_0 such that for all $n > n_0$ we have*

$$\mathbb{P}[\mathrm{EQ}(v) \geq \gamma\sqrt{d}] < \varepsilon.$$

Proof. For a vertex $v \in V_n$ we bound the value of $\mathbb{E}[\mathrm{EQ}(v)]$. Without loss of generality, we may also assume that $S_0(v) = 1$. For a vertex $w \in V_n$ to belong to $\mathrm{EQ}(v)$, we must have that $w \in N_{G(n,p)}(v)$ and $w \in \mathrm{ENC}$. Therefore, we write

$$\begin{aligned}\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 1] &= \mathbb{E}\left[\sum_{w:w\neq v} \mathbb{1}_{\{w\sim v\}}\mathbb{1}_{\{w\in \mathrm{ENC}\}} \mid S_0(v) = 1\right], \\ &= \sum_{w:w\neq v} \mathbb{P}\left[\{w \sim v\} \cap \{w \in \mathrm{ENC}\} \mid S_0(v) = 1\right].\end{aligned}$$

By conditioning on the event $\{w \sim v\}$, we have the following:

$$\begin{aligned}\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 1] &= \sum_{w:w\neq v} \mathbb{P}[w \sim v]\mathbb{P}[w \in \mathrm{ENC} \mid w \sim v, S_0(v) = 1] \\ &= \frac{d}{n}\sum_{w:w\neq v} \mathbb{P}[w \in \mathrm{ENC} \mid w \sim v, S_0(v) = 1].\end{aligned}$$

We now condition on the size of $N_{G(n,p)}(w)\backslash\{v\}$. The size of $N_{G(n,p)}(w)\backslash\{v\}$ can vary from zero to $n-2$. We set $\mu' = \mathbb{E}\left[\mathrm{EQ}(v)\right]$ and let $\mathcal{N}_k(w)$ be the event that $|N_{G(n,p)}(w)\backslash\{v\}| = k$. Note that for $w \in \mathrm{ENC}$, we necessarily have k is odd. Applying the law of total probability and bounding the upper and lower extremes of k, we have that:

$$\begin{aligned}\mathbb{P}(w \in \mathrm{ENC} \mid w \sim v, S_0(v) = 1) \leq \mathbb{P}\left(\left|\left|N_{G(n,p)}(w)\backslash\{v\}\right| - \mu'\right| \geq d^{3/4}\right) + \\ \sum_{k\;:\;|k-\mu'|\leq d^{3/4}, k\;odd} \mathbb{P}\left(|N_{G(n,p)}(w) \setminus \{v\}| = k\right)\mathbb{P}\left(w \in \mathrm{ENC}\;\middle|\{w \sim v\} \cap \mathcal{N}_k(w) \cap \{S_0(v) = 1\}\right).\end{aligned} \tag{2.26}$$

As we have conditioned on $\{v \sim w\}$, we have that $|N_{G(n,p)}(w)\backslash\{v\}| \sim \mathrm{Bin}(n-2, d/n)$. So, $\mathrm{Var}[|N_{G(n,p)}(w)\backslash\{v\}|] = (n-2)(d/n)(1-d/n)$ and therefore $\mathrm{Var}[|N_{G(n,p)}(w)\backslash\{v\}|] < d$. By Chebyshev's inequality,

$$\mathbb{P}\left(\left|\,\left|N_{G(n,p)}(w)\backslash\{v\}\right| - \mu'\right| \geq d^{3/4}\right) \leq \frac{d}{d^{6/4}} = \frac{1}{\sqrt{d}}. \tag{2.27}$$

Recall that $\mathcal{N}_k(w)$ is the event that $|N_{G(n,p)}(w)\backslash\{v\}| = k$. It follows from (2.27) that:

$$\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 1] \leq$$
$$\frac{d}{n}\sum_{w:w\neq v}\left(\frac{1}{\sqrt{d}} + \sum_{k:|k-\mu'|\leq d^{3/4},k\ odd} \mathbb{P}\left(\mathcal{N}_k(w)\right)\mathbb{P}\left(w \in \mathrm{ENC}\middle|\{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right)\right).$$

We now turn our attention to the range where $\mu - d^{3/4} \leq k \leq \mu + d^{3/4}$. We wish to apply Theorem 2.4.3 to bound $\mathbb{P}\left[w \in \mathrm{ENC} \mid \{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right]$. For a vertex $w \in N_{G(n,p)}(v)$, we define $X_w^+ = \left|\{u \in N_{G(n,p)}(w)\backslash\{v\} : S_0(u) = 1\}\right|$. As we have conditioned on $v \sim w$ and $S_0(v) = 1$, for odd k the following holds:

$$\begin{aligned}&\mathbb{P}\left(w \in \mathrm{ENC} \mid \{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right) = \\ &\qquad\mathbb{P}\left(X_w^+ = (k-1)/2 \;\middle|\; \{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right).\end{aligned} \tag{2.28}$$

Now for a given k, suppose we condition on the event $\mathcal{N}_k(w)$. By conditioning on $\mathcal{N}_k(w)$, we observe that $X_w^+ \sim \mathrm{Bin}(k, 1/2)$; so $\mathrm{Var}[X_w^+] = k/4$. We now apply Theorem 2.4.3, to bound the probability of the event given by (2.28): there exits a constant ρ such that for all k we have:

$$\begin{aligned}&\mathbb{P}\left(X_w^+ = \frac{k-1}{2}\middle| \{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right) \leq \\ &\qquad\frac{\rho}{\mathrm{Var}\left[X_k^+\right]} + \frac{1}{\sqrt{2\pi\mathrm{Var}\left[X_k^+\right]}}e^{-1/\left(8\mathrm{Var}\left[X_k^+\right]\right)} \leq \frac{2\rho}{\sqrt{\mathrm{Var}\left[X_k^+\right]}} = \frac{4\rho}{\sqrt{k}}.\end{aligned} \tag{2.29}$$

Using (2.28) and (2.29) in (2.26), we finally deduce that

$$\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 1] \leq$$
$$\frac{d(n-1)}{n}\left(\frac{1}{\sqrt{d}} + \max_{k\,:\,|k-\mu'|\leq d^{3/4},\ k\ odd}\left\{\mathbb{P}\left(X_k^+ = (k-1)/2 \;\middle|\; \{w\sim v\}\cap\mathcal{N}_k(w)\cap\{S_0(v)=1\}\right)\right\}\right).$$

Using the bound (2.29), the above expression becomes

$$\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 1] \leq \frac{d(n-1)}{n} \left(\frac{1}{\sqrt{d}} + \frac{4\rho}{\sqrt{\mu - d^{3/4}}} \right).$$

A similar argument gives the same upper bound on $\mathbb{E}[\mathrm{EQ}(v) \mid S_0(v) = 0]$. Since $\mu' = \Theta(d)$, therefore there exists a constant κ, such that for sufficiently large d we have that $\mu' - d^{3/4} \geq \kappa d$. Letting $\Gamma = 2(\sqrt{\kappa} + 4\rho)/\sqrt{\kappa}$, for sufficiently large d we have that:

$$\mathbb{E}[\mathrm{EQ}(v)] \leq \frac{\Gamma d}{\sqrt{d}} = \Gamma \sqrt{d}.$$

Fix $\varepsilon > 0$ and define $\gamma = \Gamma/\varepsilon$. Then by Markov's inequality

$$\mathbb{P}\left[\mathrm{EQ}(v) \geq \gamma \sqrt{d}\right] < \frac{\mathbb{E}[\mathrm{EQ}(v)]}{\gamma} \leq \varepsilon. \qquad \square$$

2.4.2 A Bound on the Number of Good Vertices

We now have a suitable bound on $\mathrm{EQ}(v)$, which holds with high probability. Furthermore, we would like to show that there are a significant number of good vertices. We denote the set of γ-good vertices in $G(n,p)$ as GD_γ. We recall that a vertex $v \in V_n$ of $G(n,p)$ is γ-good if v has a γ-abundant and a γ-decisive neighbourhood. The following corollary asserts that our node system will have a significant number of good vertices. We also recall the definition of the event $\mathcal{E}_c^-$: we say that the event $\mathcal{E}_c^-$ occurs if $|N_0| - |P_0| \geq 2c\sqrt{n}$, where $|P_0| = \sum_{v \in V_n} S_0(v)$ and $|N_0| = n - |P_0|$.

Corollary 2.4.5. *Let $\varepsilon > 0$ and $m = (G(n, d/n), \mathcal{S}_{1/2})^<$, where $d > \Lambda n^{1/2}$. Suppose that $\mathcal{E}_c^-$ has occurred for some constant $c > 0$. Then there exist positive constants Λ and γ, such that for all n large enough we have with probability at least $1-\varepsilon$ that $|\mathrm{GD}_\gamma| \geq n(1-\varepsilon)$.*

Proof. We fix $\varepsilon > 0$ and again define the majority game $M = (G(n, d/n), \bar{\mathcal{S}}_{1/2})^>$ as above. As we have conditioned on $\mathcal{E}_c^-$ in m, then M starts with an initial configuration which satisfies $\mathcal{E}_c^+$. By Lemma 2.4.4, there exists a constant γ such that for all n sufficiently large and all $v \in V_n$, we have that $\mathbb{P}\left[\mathrm{EQ}_\gamma(v) \geq \gamma\sqrt{d}\right] < \varepsilon^2/2$. Also, by Lemma 2.3.5, we may select Λ large enough, such that for all n sufficiently large and $v \in V_n$, we have that $\mathbb{P}\left[n_1^M(v;1) < d(v)/2 + 2\gamma\sqrt{d} \mid \mathcal{E}_c^+\right] < \varepsilon^2/2$.

Now, we denote the events that v has a γ-decisive neighbourhood by $D_\gamma(v)$ and that v has an γ-abundant neighbourhood by $A_\gamma(v)$. Note that $D_\gamma^c(v)$ and $A_\gamma^c(v)$ are the events of the previous paragraph for this particular γ. By the union bound, we have that:

$$\mathbb{P}[v \notin \mathrm{GD}_\gamma] = \mathbb{P}[A_\gamma^c(v) \cup D_\gamma^c(v)] \leq \mathbb{P}[A_\gamma^c(v)] + \mathbb{P}[D_\gamma^c(v)] < \varepsilon^2.$$

By taking a sum across all vertices, and applying the above inequality we have that $\mathbb{E}\left[|V_n \setminus \mathrm{GD}_\gamma|\right] \leq \varepsilon^2 n$, where $|V_n \setminus \mathrm{GD}_\gamma|$ is the number of vertices which are not good. By Markov's inequality, we have that $\mathbb{P}[|V_n \setminus \mathrm{GD}_\gamma| > \varepsilon n] \leq \varepsilon$. Therefore, with probability at least $1-\varepsilon$, we have that $|\mathrm{GD}_\gamma| \geq n(1-\varepsilon)$. □

Now, if there are at least $(1-\varepsilon)n$ good vertices (which occurs with high probability), then by Corollary 2.4.2 we have that $|N_2| \geq (1-\varepsilon)n$.

2.4.3 The Final Two Rounds

In the remaining two rounds we claim that unanimity will occur. We note that as a consequence of Corollary 2.4.5, at time $t = 2$ our node system satisfies the hypothesis of Lemma 2.2.16. However, as we are working within a denser regime, it turns out that unanimity can be reached in the entire $G(n,p)$ much faster than $\beta \log n$ rounds. We state a result from [32], which concerns the rapid formation of unanimity for majority dynamics, wherein one of the initial strategies has a linear majority. We recall the definitions $P_t = \{v \in V_n : S_t(v) = 1\}$, and $N_t = \{v \in V_n : S_t(v) = 0\}$. Suppose that $\mathcal{S}$ is an initial configuration of vertex states on V_n, and $\delta \in (0,1)$ is fixed. We say that $\mathcal{S} \in \hat{\mathcal{S}}_\delta$, if we have that $|N_0| \geq (1-\delta)n$ in $\mathcal{S}$. Thus $\hat{\mathcal{S}}_\delta$ is the collection of all initial vertex configurations where there are at least $(1-\delta)n$ vertices in the zero state. We now consider the following result considering the two-round evolution of games utilising initial states belonging to

$\hat{\mathcal{S}}_\delta$ for sufficiently small δ.

Lemma 2.4.6. *Let $d \geq \Lambda n^{-\frac{1}{2}}$, $p = d/n$ and $0 < \delta < 1/10$. On the probability space of $G(n,p)$ we have that a.a.s. for any $\mathcal{S} \in \mathcal{S}_\delta$ the majority game $(G(n,p), \mathcal{S})^{>}$ will reach unanimity after at most two rounds.*

While this lemma concerns the majority regime, we observe with a few adjustments to the proof, we can prove an analogous result for minority dynamics.

Lemma 2.4.7. *Let $\delta < 1/10$ and $\mathcal{S} \in S_\delta$. Then the statement of Lemma 2.4.6 holds for the minority game $(G(n,p), \mathcal{S})^{<}$.*

Proof. The proof of this lemma essentially follows the argument given in [32]. We observe that rather than the same state dominating each round, we account for the majority state switching between rounds. Suppose we have the game $m = (G(n,p), \mathcal{S})^{<}$, with $\mathcal{S} \in \mathcal{S}_\delta$ for $\delta < 1/10$. Again we recall $P_t = \{v \in V_n : S_t(v) = 1\}$ and $N_t = \{v \in V_n : S_t(v) = 0\}$. As $\mathcal{S} \in \mathcal{S}_\delta$ we have that $|N_0| \geq n(1-\delta)$. We first show a.a.s that $|N_1| < d/10$. We note that if $|N_1| \geq d/10$, then there exits a set $W \subseteq V_n$, with $|W| = d/10$ such that all vertices in W are in state zero in after round zero; we show that a.a.s such a set can not exist. For a set $W \subset V_n$ with $|W| = d/10$, let $\bar{W}$ denote the event $\{\forall v \in W : S_1(v) = 0\}$. For a set $S \subseteq V_n$ and $v \in V_n$, we recall that $d_S(v) = |N_G(v) \cap S|$. We note that,

$$\mathbb{P}(\bar{W}) \leq \mathbb{P}(\forall v \in W : d_{P_0}(v) > d_{N_0}(v)) \leq \mathbb{P}(\forall v \in W : d_{P_0}(v) \geq d_{N_0}(v)).$$

We now observe that we can decompose each of the degree terms into two parts. A degree term which intersects with W, and a degree term which intersects with W^c. By recalling that $|W| = d/10$, we have that $d_{P_0}(v) = d_{P_0 \cap W}(v) + d_{P_0/W}(v) \leq d/10 + d_{P_0/W}(v)$. Clearly we also have that $d_{N_0}(v) \geq d_{N_0/W}(v)$ therefore we have the following:

$$\mathbb{P}(\bar{W}) \leq \mathbb{P}\left(\forall v \in W : d_{P_0/W}(v) > d_{N_0/W}(v) - d/10\right).$$

This bound can alternatively be seen by observing that the most favourable position for this event to hold for v, is when all vertices of W are in the $+1$ state. We remark that for each $u, v \in W$ we have that $d_{P_0/W}(v)$ and $d_{P_0/W}(u)$ are independent, as each quantity only depends on disjoint edge sets. Therefore, $\mathbb{P}[\bar{W}]$ can be bounded by a product of event probabilities over the vertices in W. Let $v \in W$, by conditioning on whether $d_{N_0/W}(v)$ is greater than, or less than, $d/2$ and applying the law of total probability, we have:

$$\begin{aligned}\mathbb{P}\left(d_{P_0/W}(v) > d_{N_0/W}(v) - d/10\right) &\leq \mathbb{P}\left(d_{P_0/W}(v) \geq d_{N_0/W}(v) - d/10 \mid d_{N_0/W}(v) > d/10\right) \\ &\quad + \mathbb{P}\left(d_{N_0/W}(v) < d/2\right), \\ &\leq \mathbb{P}\left(d_{P_0/W}(v) \geq d/3\right) + \mathbb{P}\left(d_{N_0/w}(v) < d/2\right).\end{aligned}$$

We note that $d_{P_0/W}(v) \sim \text{Bin}(|P_0/W|, d/n)$ and $d_{N_0/W}(v) \sim \text{Bin}(|N_0/W|, d/n)$. We bound the probability of each event occurring separately. We lead with the second event, as $\mathcal{S} \in \mathcal{S}_\delta$ we have that $|N_0/W(v)| \geq n - \delta n - d/10 \geq 8n/10$ for n sufficiently large. Hence we have that $\mathbb{E}[d_{N_0/W}(v)] = d|N_0/W|/n \geq 8d/10$. Clearly we have that $|N_0/W|/n \leq 1$, hence $\mathbb{E}[d_{N_0/W}(v)] \leq d$. By applying the upper tail of Chernoff's inequality, we have:

$$\begin{aligned}\mathbb{P}(d_{N_0/W}(v) < d/2) &\leq \mathbb{P}\left(d_{N_0/w}(v) \leq (1 - 1/4)\mathbb{E}[d_{N_0/W}(v)]\right), \\ &\leq \exp(-\mathbb{E}[d_{N_0/W}(v)]/32), \\ &\leq e^{-d/32} = e^{-\Omega(d)}.\end{aligned}$$

For the other event, we note that $|P_0/W| \leq |P_0| \leq \delta n$, as $|P_0| + |N_0| = n$. Therefore, we have that $\mathbb{E}[d_{P_0/W}(v)] \leq d \cdot \delta \leq d/10$. By a similar argument we have that $\mathbb{P}(d_{P_0/W}(v) \geq d/3) = e^{-\Omega(d)}$. The above results imply that there exists a positive constant Γ, such that for every vertex $v \in W$ we have that:

$$\mathbb{P}(d_{P_0/W}(v) > d_{N_0/W}(v) - d/10) \leq e^{-\Gamma d}.$$

We note that $\bar{W}$ is the intersection on $|W|$ independent events therefore we have

$$\mathbb{P}(\bar{W}) \leq e^{-\Gamma d|W|} = e^{-\Gamma d^2/10}.$$

We take a union bound over all possible sets W with $|W| = d/10$. Hence, the probability such a W exists is bounded above by,

$$\begin{aligned}
\binom{n}{d/10} \cdot e^{-\Gamma d^2/10} \leq n^{d/10} \cdot e^{-\Gamma d^2/10} &= \exp\left(\frac{d}{10}(\log n - \Gamma d)\right), \\
&\leq \exp\left(-\frac{\Gamma d^2}{20}\right), \\
&\overset{d \geq \Lambda\sqrt{n}}{\leq} \exp\left(-\frac{\Gamma \Lambda^2 n}{20}\right).
\end{aligned}$$

We take a union bound across all partitions in $\mathcal{S}_\delta$; however, it suffices to bound across all the 2^n partitions in $\mathcal{S}$. Therefore, we have that for Λ sufficiently large,

$$\mathbb{P}(|N_1| > d/10) \leq 2^n \cdot \exp\left(-\frac{\Gamma \Lambda^2 n}{20}\right) = o(1).$$

Therefore we conclude with probability $1 - o(1)$ that $|N_1| < d/10$. In the following round we show that a.a.s $|N_2| = n$. Firstly we note that a.a.s, all vertices in G have degree at least $d/2$. Indeed, by the union bound we have that:

$$\mathbb{P}(\delta(v) \leq d/2) = \mathbb{P}\left(\bigcup_{v \in V_n} \{d_G(v) \leq d/2\}\right) \leq \sum_{v \in V_n} \mathbb{P}(d_G(v) \leq d/2).$$

We observe that $d_G(v) \sim \text{Bin}(n-1, d/n)$, therefore for sufficiently large n we have that $3d/4 \leq \mathbb{E}[d_G(v)] \leq d$. Therefore by Chernoff we have:

$$\begin{aligned}
\mathbb{P}(d_G(v) \leq d/2) &\leq \mathbb{P}\left(d_G(v) \leq \left(1 - \frac{1}{10}\right)\mathbb{E}[d_G(v)]\right), \\
&\leq \exp(-\mathbb{E}[d_G(v)]/300), \\
&\leq e^{-d/300}
\end{aligned}$$

$$\leq \quad e^{-\Lambda\sqrt{n}/300}.$$

Hence we have that $\mathbb{P}(\delta(G) \leq d/2) \leq n\exp(-\Lambda\sqrt{n}/300) = o(1)$; therefore as $|N_1| < d/10$, this implies that every vertex has a strict majority of neighbours inside P_1. Hence, all vertices switch to the zero state and thus $|N_2| = n$. As unanimity is achieved, by governing equations (2.25), it is clear that the system will also display periodic behaviour of period two. □

Proof of Theorem 2.3.2. We fix a minority game $(G, \mathcal{S})^<$, and condition on $\mathcal{E}_c^-$ for an appropriate c. By Corollary 2.4.5 and Corollary 2.4.2 there exists Λ such that for all n large enough $|N_2| > 9n/10$. Therefore this configuration now belongs to $\mathcal{S}_{1/10}$. As Lemma 2.4.7 applies to any partition in $\mathcal{S}_{1/10}$, we conclude that the system is unanimous by round four, and hence the periodic behaviour readily follows. □

2.5 Discussion

In this chapter, we studied the evolution of games on $G(n, p)$ under the best response rule. Our first main result of this chapter concerned the formation of unanimous strategies in sparser regimes, namely beyond the connectivity threshold, and on the presence of a strategic bias given by $\lambda \neq 1$. Our second result concerns the rapid formation of unanimous strategies on the node system $(G(n, p), Q, \mathcal{S}_{1/2})$ for $p \geq \Lambda n^{-1/2}$. As a byproduct of this analysis, we also prove an analogous result of Fountoulakis et al. [32] regarding the rapid formation of unanimity in the random graph minority game. A natural question of the sparse regime is to ask whether we can remove the condition that $\lambda \neq 1$. As previously seen, the case $\lambda = 1$ reduces to the problem of majority and minority dynamics. However, the study of majority dynamics for $p = o\left(n^{-1/2}\right)$ imposes immense complications. Very recently in [19], Chakraborti et al. show that Majority Dynamics stabilises to unanimity for $\lambda/n^{-1/2} \geq p \geq \lambda' n^{-3/5} \log n$. Thus combined with Theorem 2.3.1 we observe that majority dynamics achieves majority for all $p \geq \lambda' n^{-3/5} \log n$. In the most extreme case, the following was conjectured by Benjamini et al. [9].

Conjecture 2.5.1 [9]. *With high probability over the choice of random graph and choice of the initial state, if $p \geq d/n$ then the following holds:*

1. *If $d \gg 1$ then, for any $\varepsilon > 0$ and n sufficiently large we have,*

$$\lim_{t\to\infty} \Big| |P_{2t}| - |N_{2t}| \Big| \in [(1-\varepsilon)n, n] \,.$$

2. *If d is bounded then for any $\varepsilon > 0$ and for n sufficiently large,*

$$\lim_{t\to\infty} \Big| |P_{2t}| - |N_{2t}| \Big| \in [(1-\varepsilon/2)n, (1+\varepsilon/2)n].$$

These results suggest the idea of long-term almost unanimity when $d \gg 1$. Theorem 1.1.2 can be thought of as an approximate approach to studying these kinds of systems. By introducing the strategic bias $\lambda \neq 1$, we are able to show a stronger form of (1) in the above conjecture. In fact, we have been able to identify precisely the critical density of the random graph around which its largest connected component (which contains the overwhelming majority of its vertices) achieves unanimity. Furthermore, non-unanimity occurs, we identified those substructures which play different strategies to the majority of the vertices in this component. In very recent work, Charkrabroti et al. [19] consider a different range of p. Namely, they show that majority dynamics achieve unanimity in six rounds with high probability, for p such that $\lambda' n^{-3/5} \leq p \leq \lambda n^{-1/2}$, for large positive constants λ and λ'. We could hope to mimic this argument for minority dynamics, possibly with Lemma 2.4.1, this would lead to a strengthening of Theorem 1.1.1, specifically unanimity is achieved for $(G(n,p), Q, \mathcal{S})$ for $p = \Omega(n^{-3/5})$.

Another direction for consideration is to remove the synchronicity of the decision updating. For the asynchronous setting, in each round, we randomly and uniformly select a node from the network and update its strategy. Such dynamics have been considered in [7, 65]. A natural question is whether unanimity can be achieved in $G(n,p)$ under asynchronous majority dynamics, or even best response, and determine bounds on how many steps this requires. If unanimity is indeed achieved, then we can observe this must happen in at least $\Omega(n \log n)$ steps. This follows from the fact that in any initial configuration, there are a linear number of vertices in the "wrong state", and they must be sampled at some time step during the process. Due to the nature of the sampling, checking each one of these contrarian vertices follows a coupon-collector-like scheme, see [30].

CHAPTER 3

BEST RESPONSE DYNAMICS WITH MORE THAN TWO STRATEGIES

In this chapter, we study a generalisation of best response dynamics wherein vertices may choose from any of ℓ possible strategies. We recall the quantity $M(Q)$ defined in Section 1.2.2, the number of row sums in Q which achieve a maximum value. We show that unanimity occurs rapidly with high probability for the case where $M(Q) = 1$ and $M(Q) = 2$. We note that once unanimity is achieved, the system will stay unanimous and will either stabilise or cycle through some subset of the possible strategies.

3.1 Best Response with a Unique Maximum Row Sum

In this section, we consider the case where $M(Q) = 1$, and give the proof of Theorem 1.2.1. We lead this section by considering a quantity derived from the payoff matrix which will appear heavily throughout our subsequent analysis. We will also detail an expression for the difference between two entries in the score vector $T_0(v)$. We close this section with the proof of Theorem 1.2.1 and show that the system reaches unanimity in one round a.a.s.

3.1.1 Cost Coefficients

Suppose Q is a matrix then for $i, j, k, \ell \in \{0, 1, 2\}$ we define,

$$C_{i,j}^{k,\ell} = q_{i,k} - q_{i,\ell} - q_{j,k} + q_{j,\ell}.$$

We refer to these quantities as *cost coefficients*. We can consider them in the context of the minor $Q_{i,j}^{k,\ell}$ of Q, with the top left corner at (i, k) and the bottom right corner at (j, ℓ). Thus we can view the associated cost coefficient as the difference between the trace and the sum of the anti-diagonal entries of $Q_{i,j}^{k,\ell}$. As an aside, we observe that these coefficients are also proportional to the value of the zero-sum game played on $Q_{i,j}^{k,\ell}$. These quantities are central to our analysis, as they indicate the bias given by Q between pairs of strategies. We immediately observe that,

$$C_{i,j}^{\ell,k} = -C_{i,j}^{k,\ell}, \text{ and } C_{i,j}^{\ell,k} = -C_{j,i}^{\ell,k},$$

and a pair of identities that these coefficients satisfy.

Lemma 3.1.1. *Suppose for $i, j, k, \ell \in \{0, 1, 2\}$, $i \neq j$ and $k \neq \ell$ then the following identities hold:*

$$C_{i,j}^{0,1} + C_{i,j}^{1,2} = C_{i,j}^{0,2}$$

and,

$$C_{0,1}^{k,\ell} + C_{1,2}^{k,\ell} = C_{0,2}^{k,\ell}.$$

Proof. Both equalities can be immediately verified by direct computation. We illustrate the former,

$$\begin{aligned} C_{i,j}^{0,1} + C_{i,j}^{1,2} &= (q_{i,0} - q_{i,1} - q_{j,0} + q_{j,1}) + (q_{i,1} - q_{i,2} - q_{j,1} + q_{j,2}) \\ &= q_{i,0} - q_{j,0} - q_{i,2} + q_{j,2} \\ &= C_{i,j}^{0,2}. \end{aligned}$$

□

3.1.2 Decomposition of the Score Difference and Proof of Theorem 1.2.1

We detail a general expression for the comparison of scores between strategies i and j, namely the term $T_0(v;i) - T_0(v;j)$, we refer to this term as the *payoff difference* between i and j. We derive a decomposition of the payoff difference in terms of a leading term, of order $O(np)$, and an additional deviation term of lower order. For each $i \in \{0, \ldots, \ell - 1\}$ we denote the quantity $d_0(v;i)$ to be such that,

$$n_0(v;i) = \frac{|N_G(v)|}{\ell} + d_0(v;i).$$

Lemma 3.1.2. *Let* $\ell \geq 2$ *be an integer. Then we have for every* $0 \leq i < j \leq \ell - 1$:

$$\mathbb{P}\left[T_0(v;i) > T_0(v;j)\right] = \mathbb{P}\left[\frac{|N_G(v)|}{\ell}(\Sigma R_i - \Sigma R_j) + \sum_{k=0}^{\ell-2} C_{i,j}^{k,\ell-1} d_0(v;k) > 0\right].$$

Proof. We lead by considering the difference in scores between strategy i and strategy j :

$$\begin{aligned}
T_0(v;i) - T_0(v;j) &= \sum_{k=0}^{\ell-1} n_0(v;k)(q_{i,k} - q_{j,k}) \\
&= \frac{|N_G(v)|}{\ell} \sum_{k=0}^{\ell-1}(q_{i,k} - q_{j,k}) + \sum_{k=0}^{\ell-1} d_0(v;k)(q_{i,k} - q_{j,k}).
\end{aligned}$$

We recall for $k \in \{0, 1, 2\}$ we denote ΣR_k to be the sum of elements in row k of the matrix Q. Hence, we observe that the first summation is the difference of ΣR_i and ΣR_j For the second summation, we observe that

$$\begin{aligned}\sum_{k=0}^{\ell-1} d_0(v;k) &= \sum_{k=0}^{\ell-1} n_0(v;k) - \frac{|N_G(v)|}{\ell} \\ &= |N_G(v)| - |N_G(v)| \\ &= 0.\end{aligned}$$

Hence we can express the final deviation in terms of the former $\ell - 2$ deviations,

$$\begin{aligned}T_0(v;i) - T_0(v;j) &= \frac{|N_G(v)|}{\ell}(\Sigma R_i - \Sigma R_j) + \sum_{k=0}^{\ell-2} d_0(v;k)\left(q_{i,k} - q_{j,k} - (q_{i,\ell-1} - q_{j,\ell-1})\right) \\ &= \frac{|N_G(v)|}{\ell}(\Sigma R_i - \Sigma R_j) + \sum_{k=0}^{\ell-2} d_0(v;k) C_{i,j}^{k,\ell-1}.\end{aligned}$$ □

The condition that Q has a unique maximum row sum is sufficient for determining the initial strategy first achieved at unanimity. For any given i and j, we have that $T_0(v;i) - T_0(v;j)$ has a leading term whose sign depends on the difference of the row sums. To this end, we declare unanimity in one round to the strategy whose row is the maximal row sum in Q. We re-state the result as follows:

Theorem 1.2.1. *Let p be such that, $p \gg n^{-\alpha}$, for $\alpha \in (0,1)$. Suppose $\ell \in \mathbb{N}$ and Q is an $\ell \times \ell$ payoff matrix. Furthermore suppose $M(Q) = 1$, $i^* = \text{argmax}_{0 \le i \le \ell-1}\{\Sigma R_i\}$, and the column maxima of Q are uniquely defined. Then across the product space of $\mathcal{S}_{1/\ell}$ and $G = G(n,p)$ we have with probability $1 - o(1)$ that $S_1(v) = i^*$ for every $v \in V_n$.*

Proof. We lead by considering the comparison of scores between strategy i^* and an arbitrary strategy j. We have by Lemma 3.1.2

$$\mathbb{P}\left[T_0(v;i^*) > T_0(v;j)\right] = \mathbb{P}\left[\frac{|N_G(v)|}{\ell}(\Sigma R_{i^*} - \Sigma R_j) + \sum_{k=0}^{\ell-2} C_{i,j}^{k,\ell-1} d_0(v;k) > 0\right], \tag{3.1}$$

with $d_0(v;k) = n_0(v;k) - |N_G(v)|/\ell$. Thus as ΣR_{i^*} is the unique maximum row sum we have that $\Sigma R_{i^*} > \Sigma R_j$ for every $j \neq i^*$. Thus for the range of p we are considering the leading term on the right hand side of (3.1) will be positive and have order $O(np)$. While we will show the remaining terms have order at most $O(\sqrt{np}\log n)$ and thus we have that $T_0(v;i^*) > T_0(v;j)$ for every j. Firstly fix $v \in V(G)$ then by the law of total probability we can condition on the size of $|N_G(v)|$ as follows:

$$\mathbb{P}[T_0(v;i^*) < T_0(v;j)] = \sum_{t=0}^{n-1} \mathbb{P}\left[T_0(v;i^*) < T_0(v;j) \middle| |N_G(v)| = t\right] \mathbb{P}[|N_G(v)| = t]$$

We split the summation into two parts depending on the value of $|N_G(v)|$. For $|N_G(v)|$ further than $\sqrt{np}\log n$ away from it's expectation, we can bound this probability using Chernoff. Hence we have that $\mathbb{P}[T_0(v;i^*) < T_0(v;j)]$ is bounded above by,

$$2\sqrt{np}\log n \max_{t:|t-np|\leq\sqrt{np}\log n} \left\{ \mathbb{P}\left[\frac{|N_G(v)|}{\ell}(R_{i^*} - R_j) + \sum_{k=0}^{\ell-2} d_0(v;k) C_{i^*,j}^{k,\ell-1} < 0 \middle| |N_G(v)| = t\right]\right\}$$
$$O(e^{-(\log n)^2}).$$

The third inequality holds due to a standard Chernoff bound on the concentration of the vertex degrees, we consider this more precisely in Section 3.2.1. We have that $R_{i^*} > R_j$, hence as we have conditioned on t being the range of $np \pm \sqrt{np}\log n$ it follows there exists a constant $K > 0$, such that for t in this range,

$$\begin{aligned}
&\mathbb{P}\left[\frac{|N_G(v)|}{\ell}(\Sigma R_{i^*} - \Sigma R_j) + \sum_{k=0}^{\ell-2} d_0(v;k) C_{i^*,j}^{k,\ell-1} < 0 \middle| |N_G(v)| = t\right] \\
&\leq \mathbb{P}\left[\left|\sum_{k=0}^{\ell-2} d_0(v;k) C_{i^*,j}^{k,\ell-1}\right| \geq \frac{1}{2}(np)^{1/2}\log n \middle| |N_G(v)| = t\right]
\end{aligned}$$

Hence if the sum of a constant number of these terms terms is larger than $\frac{1}{2}(np)^{1/2}\log n$, we must have that at least one of them of is of at least this order. Therefore, with the union we bound, we have

$$\mathbb{P}\left[\bigcup_{k=0}^{\ell-2}\{|d_0(v;k)| \geq \frac{1}{2K}(np)^{1/2}\log n\} \middle| |N_G(v)| = t\right]$$

$$\leq \sum_{k=0}^{\ell-2} \mathbb{P}\left[\{|d_0(v;k)| \geq \frac{1}{2K}(np)^{1/2}\log n\}\middle| |N_G(v)| = t\right].$$

Thus it suffices for us to the bound the probability of the above event uniformly over k. We note that conditional on $|N_G(v)| = t$, we have that $n_0(v;k) \sim \text{Bin}(t, 1/\ell)$ hence, $\mathbb{E}(n_0(v;k)) = t/\ell$. Thus we have that $d_0(v;k) = n_0(v;k) - \mathbb{E}[n_0(v;k)]$. Thus by Chernoff we have:

$$\mathbb{P}\left[|d_0(v;k)| \geq \frac{1}{2K}(np)^{1/2}\log n\middle| |N_G(v)| = t\right] = \mathbb{P}\left[|n_0(v;k) - \mathbb{E}[n_0(v;k)]| \geq \frac{1}{2K}(np)^{1/2}\log n\right]$$
$$\leq e^{-\frac{\ell(np)\log^2 n}{4Kt}}.$$

Thus substituting this into the above expression, we have for a vertex v that:

$$\mathbb{P}[T_0(v;i^*) < T_0(v;j)] \leq 2\sqrt{np}\log n \max_{t:|t-np|\leq\sqrt{np}} \left\{e^{-\frac{\ell(np)\log^2 n}{4Kt}}\right\} + O\left(e^{-(\log n)^2}\right)$$
$$\leq 3\sqrt{np}\log n \cdot e^{-\Omega(\log^2 n)} + O\left(e^{-(\log n)^2}\right).$$

Finally, we take a union bound across every strategy compared with i^* and every vertex, thus we have:

$$\mathbb{P}\left[\bigcup_{v\in V(G)} \{S_1(v) \neq i^*\}\right] \leq \mathbb{P}\left[\bigcup_{v\in V(G)} \bigcup_{\substack{0\leq j\leq \ell-1\\ j\neq i^*}} \{T_0(v;i^*) < T_0(v;j)\}\right]$$
$$\leq \sum_{v\in V(G)} \sum_{\substack{0\leq j\leq \ell-1\\ j\neq i^*}} \mathbb{P}\left[\{T_0(v;i^*) < T_0(v;j)\}\right].$$
$$\leq 3n(\ell-1)\sqrt{np}\log n \cdot e^{-\Omega(\log^2 n)} + 3n(\ell-1)\cdot O\left(e^{-\log^2 n}\right).$$
$$= o(1).$$

□

We highlight two direct corollaries of Theorem 1.2.1. Firstly, for any system as above, with $\Sigma R_{i^*} > \Sigma R_j$, then all vertices will never switch to strategy j a.a.s. We summarise this as follows.

Corollary 3.1.3. *Suppose $p \gg n^{-\alpha}$ and Q is an $\ell \times \ell$ payoff matrix as above. Suppose* *that $\Sigma R_i^* > \Sigma R_j$, then with probability at least $1 - o(1)$ that for all $v \in V_n$ we have,*

$$T_0(v; i^*) > T_0(v; j).$$

Our second corollary states that in the case of the 3×3 matrix game, for all $v \in V_n$, the initial deviations of the zero and one strategies are at most $\log n(np)^{1/2}$ a.a.s.

Corollary 3.1.4. *With probability $1-o(1)$ any vertex $v \in V_n$ has $|d_0(v;0)| \leq \log n \cdot (np)^{1/2}$ and $|d_0(v;1)| \leq \log n \cdot (np)^{1/2}$.*

3.2 Best Response for $M(Q) = 2$

In this section, we proceed with the proof of Theorem 1.2.2. Namely for 3×3 matrix Q with exactly two equal maximal row sums, we show that the system reaches unanimity in three steps with high probability. We restate our main result as follows,

Theorem 1.2.2. *Let $np \gg n^{2/3}$. Suppose that Q is a 3×3 payoff matrix, with $M(Q) = 2$ and a unique column maxima. Then for every $\varepsilon > 0$ there exists an $n_0 \in \mathbb{N}$ such that for all $n > n_0$, across the product space of $G(n,p)$ and $\mathcal{S}_{1/3}$, we have that there exists some $i^* \in \{0, 1, 2\}$ such that*

$$\mathbb{P}\left[\bigcap_{v \in V_n} \{S_3(v) = i^*\}\right] \geq 1 - \varepsilon.$$

We lead this section with the derivation of a Local Limit Theorem for the number of strategies of each type in the neighbourhood of v, from the perspective of the hypergeometric distribution. We then consider a decomposition of the payoff difference, in a similar vein to Lemma 3.1.2. Consequently, we determine a condition for either of the two-row dominating strategies to gain a bias. By applying our local limit theorem, we estimate these probabilities and find bounds on the size of this bias. Following this, we show another strategy gains a significant bias in the next round, followed by the declaration of unanimity after the third round with high probability.

3.2.1 A Local Limit Theorem

Consider a vertex $v \in V_{n+1}$ and a set $\mathcal{S}_n = V_{n+1} \setminus \{v\}$ of size n[1] We consider the random neighbourhood $N(v)$ of v in $\mathcal{S}_n$, where each vertex in $\mathcal{S}_n$ becomes a neighbour of v, with probability p independently of any other vertex in V_{n+1}. Let $n(v)$ denote the size of its neighbourhood in $\mathcal{S}_n$, which follows the binomial distribution $\text{Bin}(n, p)$. Let us condition on the event that $|n(v) - np| \leq \log n \cdot (np)^{1/2}$ - we call this $\mathcal{N}_v$. The standard Chernoff bound yields that $\mathcal{N}_v$ occurs with probability $1 - o(1/n)$. In general, these parts represent the set of vertices playing each strategy. Throughout this section, $\mathcal{S}_n$ is partitioned into three parts $\{S_0(0), S_0(1), S_0(2)\}$, where for each $v \in \mathcal{S}_n$ we have $v \in S_0(i)$ if $S_0(v) = i$. We aim to quantify the bias given to each strategy from the initial multinomial global distribution $\mathcal{S}_{1/3}$. To account for this, we define c_i such that,

$$|S_0(i)| = n_i \text{ and } n_i = n/3 + c_i\sqrt{n}. \tag{3.2}$$

We will justify this choice of bias from $n/3$ in Section 3.2.2.

Let us write $n_i = np_i$; thus $p_i = 1/3 + c_i/\sqrt{n}$. Let us condition on $\mathcal{N}_v$ and, in particular, the value of $n(v)$ is such that $\mathcal{N}_v$ is satisfied. For $i = 0, 1$ let us write $m_i = \lceil n(v)p_i \rceil$ and $m_2 = n(v) - m_0 - m_1$. Let $n(v; i) = |N(v) \cap S_i|$, for $i = 0, 1, 2$. We will give an estimate of the probability that $(n(v; 0), n(v; 1), n(v; 2)) = (k_0, k_1, k_2)$ conditional on the value of $n(v)$. Note that this probability is:

$$\mathbb{P}\left[(n(v; 0), n(v; 1), n(v; 2)) = (k_0, k_1, k_2)\right] = \frac{\binom{n_0}{k_0} \cdot \binom{n_1}{k_1} \cdot \binom{n_2}{k_2}}{\binom{n}{n(v)}}. \tag{3.3}$$

We will show that this probability is approximated by the probability distribution function of a bivariate gaussian random vector, provided that k_i is close to m_i for $i = 0, 1$. We remark that such a probability is dependent on the revealing of $n(v)$ and the initial global strategies $\mathcal{S}_n = \{S_0(0), S_0(1), S_0(2)\}$. With a slight abuse of notation, for an event $\mathcal{E}$ we

[1] Note that when we apply the following calculations, we will work with the set S_{n-1} as G is an n vertex graph. This choice allows a sequence of slightly neater calculations.

write,

$$\mathbb{P}\left[\mathcal{E}|n(v),\mathcal{S}_n\right] = \mathbb{E}\left(\mathbb{1}_{\mathcal{E}}|n(v),\mathcal{S}_n\right).$$

Lemma 3.2.1. *Let $p = p(n) \in [0,1]$ be such that $\limsup_{n\to\infty} p(n) < 1$ and $n^{1/2}p \geq \log^3 n$. Let $n(v)$ be such that $\mathcal{N}_v$ occurs. Suppose that $k_i = m_i + \delta_i$ with $|\delta_i| \leq \log n \cdot (np)^{1/2}$, for $i = 0,1$. Then uniformly over these choices of (δ_0,δ_1) with $\mathbf{x} = (x_0,x_1)^T := (\delta_0,\delta_1)^T/n(v)^{1/2}$ and*

$$\Sigma = \frac{1}{1-n(v)/n}\begin{bmatrix} \frac{1}{p_0}+\frac{1}{p_2} & \frac{1}{p_2} \\ \frac{1}{p_2} & \frac{1}{p_1}+\frac{1}{p_2} \end{bmatrix},$$

we have

$$\mathbb{P}\left[(n(v;0),n(v;1),n(v;2)) = (k_0,k_1,k_2)|n(v),\mathcal{S}_n\right] = \frac{1}{n(v)}\cdot\phi(\mathbf{x})\left(1+o(p^{1/2})\right),$$

where $\phi(\mathbf{x}) = \frac{1}{2\pi}|\Sigma|^{1/2}\exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma\mathbf{x}\right)$.

Proof. We will give asymptotic estimates for each one of the four binomial coefficients that appear in (3.3). Before we do this, let us give some relations, which we will use in our estimates later. Firstly, let us observe

$$\frac{1}{np} \leq \frac{\log^3 n}{\sqrt{np}} = o\left(p^{1/2}\right). \tag{3.4}$$

This is equivalent to $1/(np)^2 = o(p)$ which in turn is equivalent to $1/n^2 = o(p^3)$. But this holds, since $1/n^{1/2} = o(p)$. Also,

$$\frac{\log^2 n}{np} = o\left(p^{1/2}\right). \tag{3.5}$$

This is equivalent to $\log^2 n/n = o(p^{3/2})$. But

$$\frac{\log^{4/3} n}{n^{2/3}} \ll \frac{1}{n^{1/2}} \ll p.$$

We begin by estimating the binomial coefficient that appears in the denominator of (3.3). Let

$$H(x) = -x \ln x - (1-x)\ln(1-x),$$

denote the entropy function defined for $x \in (0,1)$. We recall a standard estimate for this binomial coefficient which relies on the Stirling approximation for the factorial: $n! = \sqrt{2\pi n} n^n e^{-n}(1 + O(1/n))$. Using this, we write

$$\begin{aligned}
\binom{n}{n(v)} &= \frac{n!}{n(v)!(n-n(v))!} \\
&= \frac{1+O(1/n(v))}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{n(v)(n-n(v))}} \cdot \frac{n^n e^{-n}}{n(v)^{n(v)} e^{-n(v)} (n-n(v))^{n-n(v)} e^{-(n-n(v))}} \\
&= \frac{1+O(1/np)}{\sqrt{2\pi}} \cdot \sqrt{\frac{n}{n(v)(n-n(v))}} e^{nH(n(v)/n)}.
\end{aligned}$$

Therefore,

$$\binom{n}{n(v)} = (1 + o(p^{1/2}))\sqrt{\frac{1}{1-n(v)/n}} \cdot \frac{1}{\sqrt{2\pi n(v)}} e^{nH(n(v)/n)}. \tag{3.6}$$

Now, we will consider $\binom{n_i}{k_i}$. In fact, we will express this in terms of $\binom{n_i}{m_i}$. In turn, we will express the latter using (3.6). Recalling that $k_i = m_i + \delta_i$, we write

$$\binom{n_i}{k_i} \Big/ \binom{n_i}{m_i} = \frac{m_i!}{(m_i+\delta_i)!} \cdot \frac{(n_i - m_i)!}{(n_i - m_i - \delta_i)!}.$$

Suppose first that $\delta_i > 0$. Then

$$\frac{m_i!}{(m_i+\delta_i)!} = \frac{1}{m_i^{\delta_i}} \prod_{j=1}^{\delta_i} \left(1 + \frac{j}{m_i}\right)^{-1}.$$

Writing $1 + j/m_i = \exp\left(\ln(1 + j/m_i)\right)$ and expanding the ln function around 1, we get that

$$\prod_{j=1}^{\delta_i} \left(1 + \frac{j}{m_i}\right)^{-1} = \exp\left(-\frac{\delta_i^2}{2m_i} + O\left(\frac{|\delta_i|}{m_i} + \frac{|\delta_i|^3}{m_i^2}\right)\right).$$

Similarly, we get

$$\frac{(n_i - m_i)!}{(n_i - m_i - \delta_i)!} = (n_i - m_i)^{\delta_i} \exp\left(-\frac{\delta_i^2}{2(n_i - m_i)} + O\left(\frac{|\delta_i|}{n_i - m_i} + \frac{|\delta_i|^3}{(n_i - m_i)^2}\right)\right).$$

As $n_i - m_i = \Omega(n)$, we deduce that

$$\frac{|\delta_i|}{n_i - m_i} + \frac{|\delta_i|^3}{(n_i - m_i)^2} = O\left(\frac{\log n \cdot (np)^{1/2}}{n} + \frac{\log^3 n \cdot (np)^{3/2}}{n^2}\right) = o(p^{1/2}).$$

We thus conclude that

$$\binom{n_i}{k_i} = \left(1 + o(p^{1/2}) + O\left(\frac{|\delta_i|}{np} + \frac{|\delta_i|^3}{(np)^2}\right)\right)\binom{n_i}{m_i} \cdot \left(\frac{n_i - m_i}{m_i}\right)^{\delta_i} \cdot \exp\left(-\frac{\delta_i^2}{2} \cdot \frac{n_i}{m_i(n_i - m_i)}\right). \tag{3.7}$$

Note that as $|n(v) - np| \leq (np)^{1/2} \log n$, we have

$$\frac{m_i}{n_i} = \frac{n(v)}{n} + O(1/n) \tag{3.8}$$

and

$$m_i = n(v)p_i(1 + O(1/np)) \overset{(3.4)}{=} n(v)p_i(1 + o(p^{1/2})). \tag{3.9}$$

Since $|\delta_i| = O(\log n \cdot (np)^{1/2})$, we deduce that

$$\left(\frac{n_i - m_i}{m_i}\right)^{\delta_i} = \left(\frac{n}{n(v)} - 1\right)^{\delta_i}\left(1 + O(\log n\sqrt{p/n})\right) = \left(\frac{n}{n(v)} - 1\right)^{\delta_i}(1 + o(p^{1/2})).$$

Furthermore by (3.7) and (3.8) we have,

$$\frac{\delta_i^2}{2} \cdot \frac{n_i}{m_i(n_i - m_i)} = \frac{\delta_i^2}{2n(v)p_i}\left(\frac{1}{1 - n(v)/n}\right) + O(\delta_i^2/(np)^2).$$

But $\delta_i^2 = O(\log^2 n \cdot (np))$ and therefore

$$\frac{\delta_i^2}{2} \cdot \frac{n_i}{m_i(n_i - m_i)} = \frac{\delta_i^2}{2n(v)p_i}\left(\frac{1}{1 - n(v)/n}\right) + O(\log^2 n/(np)).$$

But $\log^2 n/(np) \overset{(3.5)}{=} o(p^{1/2})$ and thereby

$$\exp\left(-\frac{\delta_i^2}{2} \cdot \frac{n_i}{m_i(n_i - m_i)}\right) = (1 + o(p^{1/2}))\exp\left(-\frac{\delta_i^2}{2n(v)p_i}\left(\frac{1}{1 - n(v)/n}\right)\right).$$

As $n_i = \Theta(n)$ and $m_i = \Theta(np)$, applying (3.6), we can also deduce that

$$\binom{n_i}{m_i} = (1 + o(p^{1/2}))\sqrt{\frac{1}{1 - n(v)/n}} \cdot \frac{1}{\sqrt{2\pi m_i}} e^{n_i H(m_i/n_i)}. \tag{3.10}$$

Combining (3.6), (3.7) and (3.10), we deduce that

$$\mathbb{P}\left[(n(v;0), n(v;1), n(v;2)) = (k_0, k_1, k_2)\right] =$$

$$\left(1 + o(p^{1/2}) + O\left(\frac{|\delta_0| + |\delta_1|}{np} + \frac{|\delta_0|^3 + |\delta_1|^3}{(np)^2}\right)\right) \times \frac{1}{1 - n(v)/n} \cdot \frac{1}{2\pi} \cdot \sqrt{\frac{n(v)}{m_0 m_1 m_2}} \cdot \left(\prod_{i=0}^{2}\left(\frac{n}{n(v)} - 1\right)^{\delta_i}\right)$$
$$\times \exp\left(-\frac{1}{2}\sum_{i=0}^{2} \frac{\delta_i^2}{n(v)} \frac{1}{p_i(1 - n(v)/n)}\right) \cdot e^{\sum_{i=0}^{2} n_i H(m_i/n_i) - nH(n(v)/n)}.$$

As we have that $\delta_0 + \delta_1 + \delta_2 = 0$ it follows that the above is equal to,

$$\left(1 + o(p^{1/2}) + O\left(\frac{|\delta_0| + |\delta_1|}{np} + \frac{|\delta_0|^3 + |\delta_1|^3}{(np)^2}\right)\right) \times \frac{1}{1 - n(v)/n} \cdot \frac{1}{2\pi} \cdot \frac{1}{n(v)} \cdot \sqrt{\frac{1}{p_0 p_1 p_2}}$$
$$\times \exp\left(-\frac{1}{2}\sum_{i=0}^{2} \frac{\delta_i^2}{n(v)} \frac{1}{p_i(1 - n(v)/n)}\right) \times e^{\sum_{i=0}^{2} n_i H(m_i/n_i) - nH(n(v)/n)}. \tag{3.11}$$

Claim 3.2.2. *With* $\mathbf{x} = (\delta_0, \delta_1)^T / n(v)^{1/2}$ *and,*

$$\Sigma = \frac{1}{1 - n(v)/n}\begin{bmatrix} \frac{1}{p_0} + \frac{1}{p_2} & \frac{1}{p_2} \\ \frac{1}{p_2} & \frac{1}{p_1} + \frac{1}{p_2} \end{bmatrix}$$

we have

$$\exp\left(-\frac{1}{2}\sum_{i=0}^{2} \frac{\delta_i^2}{n(v)} \frac{1}{p_i(1 - n(v)/n)}\right) = \exp\left(-\frac{1}{2}\mathbf{x}^T \Sigma \mathbf{x}\right).$$

Proof of Claim. Since $\delta_0 + \delta_1 + \delta_2 = 0$ we write with $\delta = (\delta_0, \delta_1)^T$

$$\sum_{i=0}^{2} \frac{\delta_i^2}{p_i} = \frac{\delta_0^2}{p_0} + \frac{\delta_1^2}{p_1} + \frac{(\delta_0 + \delta_1)^2}{p_2}$$

$$= \delta_0^2\left(\frac{1}{p_0}+\frac{1}{p_2}\right)+\delta_1^2\left(\frac{1}{p_1}+\frac{1}{p_2}\right)+\delta_0\delta_1\frac{2}{p_2}$$
$$= \delta^T\begin{bmatrix} \frac{1}{p_0}+\frac{1}{p_2} & \frac{1}{p_2} \\ \frac{1}{p_2} & \frac{1}{p_1}+\frac{1}{p_2} \end{bmatrix}\delta.$$

Thus, setting

$$\Sigma = \frac{1}{1-n(v)/n}\begin{bmatrix} \frac{1}{p_0}+\frac{1}{p_2} & \frac{1}{p_2} \\ \frac{1}{p_2} & \frac{1}{p_1}+\frac{1}{p_2} \end{bmatrix},$$

and using the scaling $\mathbf{x} = (\delta_0, \delta_1)^T/n(v)^{1/2}$ we get

$$\exp\left(-\frac{1}{2}\sum_{i=0}^{2}\frac{\delta_i^2}{n(v)}\frac{1}{p_i(1-n(v)/n)}\right) = \exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma\mathbf{x}\right).$$

$\diamondsuit$

We now calculate $|\Sigma|$.

Claim 3.2.3. *We have*

$$|\Sigma| = \frac{1}{(1-n(v)/n)^2}\cdot\frac{1}{p_0p_1p_2}.$$

Proof of Claim. We have

$$|\Sigma| = \frac{1}{(1-n(v)/n)^2}\cdot\left(\left(\frac{1}{p_0}+\frac{1}{p_2}\right)\left(\frac{1}{p_1}+\frac{1}{p_2}\right)-\frac{1}{p_2^2}\right).$$

But as $p_0+p_1+p_2 = 1$ we have,

$$\left(\frac{1}{p_0}+\frac{1}{p_2}\right)\left(\frac{1}{p_1}+\frac{1}{p_2}\right)-\frac{1}{p_2^2} = \frac{1}{p_0p_1p_2}.$$

We thus conclude that

$$|\Sigma| = \frac{1}{(1-n(v)/n)^2}\cdot\frac{1}{p_0p_1p_2}.$$

$\diamondsuit$

Now, we will deal with the last exponential in (3.11).

Claim 3.2.4. *We have*

$$\sum_{i=0}^{2} n_i H(m_i/n_i) - nH(n(v)/n) = O(1/np).$$

Proof of Claim. We will approximate $H(m_i/n_i)$ by $H(n(v)/n)$ using (the second order) Taylor's theorem around $n(v)/n$. Let $i \in \{0,1\}$. Then $m_i \geq n(v)p_i$, whereby $n(v)/n \leq m_i/n_i$. Furthermore, for n sufficiently large $m_i/n_i < 1$ and therefore $[n(v)/n, m_i/n_i] \subset (0,1)$. Since the entropy function H is twice differentiable in $(0,1)$, there exists $\xi \in [n(v)/n, m_i/n_i]$

$$H\left(\frac{m_i}{n_i}\right) = H\left(\frac{n(v)}{n_i}\right) + H'\left(\frac{n(v)}{n}\right)\left(\frac{m_i n - n_i n(v)}{nn_i}\right) + \frac{1}{2}H''(\xi)\left(\frac{m_i n - n_i n(v)}{nn_i}\right)^2.$$

By (3.8), we have that

$$\left(\frac{m_i n - n_i n(v)}{nn_i}\right)^2 = O\left(\frac{1}{n^2}\right).$$

Also, $H''(x) = -1/(x - x^2)$, for any $x \in (0,1)$. As $\xi = p + o(p)$ and p is asymptotically bounded away from 1, we have

$$|H''(\xi)| = \frac{1}{\xi(1-\xi)} = O(1/\xi) = O(1/p).$$

We can deduce the same for $i = 2$, except that in that case $m_2/n_2 \leq n(v)/n$. Using these, we can write

$$\begin{aligned}
\sum_{i=0}^{2} n_i H\left(\frac{m_i}{n_i}\right) &= \sum_{i=0}^{2} n_i H\left(\frac{n(v)}{n}\right) + \sum_{i=0}^{2} n_i H'\left(\frac{n(v)}{n}\right)\left(\frac{m_i n - n_i n(v)}{nn_i}\right) + O\left(\frac{1}{n^2 p}\right)\sum_{i=0}^{2} n_i \\
&= nH\left(\frac{n(v)}{n}\right) + \sum_{i=0}^{2} H'\left(\frac{n(v)}{n}\right)\left(\frac{n\sum_i m_i - n(v)\sum_i n_i}{n}\right) + O\left(\frac{1}{np}\right) \\
&= nH\left(\frac{n(v)}{n}\right) + O\left(\frac{1}{np}\right),
\end{aligned}$$

since $\sum_i m_i = n(v)$ and $\sum_i n_i = n$. ◊

Therefore,

$$e^{\sum_{i=0}^{2} n_i H(m_i/n_i) - nH(n(v)/n)} = e^{O(1/np)} = 1 + O(1/(np)) \stackrel{(3.4)}{=} 1 + o(p^{1/2}). \tag{3.12}$$

Finally we observe that as $|\delta_i| \leq (np)^{1/2} \log n$,

$$\frac{|\delta_0| + |\delta_1|}{np} + \frac{|\delta_0|^3 + |\delta_1|^3}{(np)^2} = O\left(\frac{\log^3 n}{\sqrt{np}}\right) \stackrel{(3.4)}{=} o(p^{1/2}).$$

We thus conclude that

$$\mathbb{P}\left[(n(v;0), n(v;1), n(v;2)) = (k_0, k_1, k_2) | n(v), \mathcal{S}_n\right] = \frac{1}{n(v)} \cdot \phi(\mathbf{x}) \cdot \left(1 + o(p^{1/2})\right) \qquad \square$$

3.2.2 Some (Anti-)Concentration Properties of the Initial Configuration

Let $\mathcal{S}_0 : V_n \to \{0,1,2\}$ be the random configuration on V_n in which $S_0(v) = i \in \{0,1,2\}$ with probability 1/3, independently of any other vertex in V_n. We recall that $S_0(i) \subseteq V_n$ denotes the subset of V_n which consists of those vertices in V_n which play strategy i in the initial round. Then $|S_0(i)| \sim \mathrm{Bin}(n, 1/3)$. Let us also recall the local limit theorem for the binomial distribution.

Theorem 3.2.5. *Let $X \sim \mathrm{Bin}(n,q)$ for some fixed $q \in (0,1)$. Then with $\sigma^2 = nq(1-q)$ we have*

$$\sup_{k \in \mathbb{Z}} \left| \mathbb{P}\left[X = k\right] - \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(k-np)^2}{2\sigma^2}} \right| = O\left(\frac{1}{nq}\right).$$

The local limit theorem implies that the probability that $|S_0(i)|$ is too close to $n/3$ (applying it for $q = 1/3$) is small.

Lemma 3.2.6. *For any $\delta > 0$ there exists $\varepsilon > 0$ such that for n sufficiently large,*

$$\mathbb{P}\left[\left||S_0(i)| - n/3\right| \leq \varepsilon\sqrt{n}\right] < \delta.$$

Proof. We use the local limit theorem above and write

$$\begin{aligned}
\mathbb{P}\left[|S_0(i)| - n/3| \le \varepsilon\sqrt{n}\right] &= \sum_{k:|k-n/3|\le\varepsilon\sqrt{n}} \mathbb{P}\left[|S_0(i)| = k\right] \\
&= \frac{1}{\sqrt{2\pi n(1/3)(2/3)}} \sum_{k:|k-n/3|\le\varepsilon\sqrt{n}} e^{-\frac{(k-n/3)^2}{2n(1/3)(2/3)}} + O\left(\frac{1}{\sqrt{n}}\right) \\
&\le \frac{2\varepsilon\sqrt{n}}{\sqrt{2\pi n(1/3)(2/3)}} + O\left(\frac{1}{\sqrt{n}}\right) = O(\varepsilon).
\end{aligned}$$

The lemma follows. □

Also, the Chernoff bound implies that $|S_0(i)|$ is not too far from $n/3$.

Lemma 3.2.7. *For any $\delta > 0$, there exists $C > 0$ such that*

$$\mathbb{P}\left[|S_0(i)| - n/3| > C\sqrt{n}\right] < \delta.$$

The two lemmas together with the union bound imply that for any δ there exists $\varepsilon > 0$ such that with probability at least $1 - \delta$, for $i = 0, 1, 2$ we have

$$\varepsilon < \frac{1}{\sqrt{n}} ||S_0(i)| - n/3| \le 1/\varepsilon.$$

We denote this event $\mathcal{E}_{n,\varepsilon}$. In what follows, we will condition on $\mathcal{E}_{n,\varepsilon}$. More specifically, we shall assume that the initial configuration $\mathcal{S}_0$ induces the partition $(S_0(0), S_0(1), S_0(2))$ where

$$|S_0(i)| = n/3 + c_i\sqrt{n},$$

where $\varepsilon < |c_i| < 1/\varepsilon$. For such initial states satisfying the above, we write $\mathcal{S}_0 \in \mathcal{E}_{n,\varepsilon}$. Note that $c_0 + c_1 + c_2 = 0$.

3.2.3 The Expected Evolution of the System After One Step

Let $v \in V_n$. We will estimate the probability that $S_1(v) = i$, for $i \in \{0, 1, 2\}$. Without loss of generality, let us assume that rows 0 and 1 have maximal row sums, i.e. both strictly larger than ΣR_2. Previously we have shown that the probability that $S_1(v) = 2$ is $o(1/n)$, this is given by Theorem 1.2.1 and Corollary 3.1.3, hence the probability a vertex has $S_1(v) = 2$ is equal to $o(1)$. For the remainder of this section, we condition on this event for all $v \in V_n$. We show that the probabilities of the events $S_1(v) = 0$, $S_1(v) = 1$ are approximately 1/2. However, we will show that one of them is larger than 1/2 by a term that is of order $p^{1/2}$. The chosen strategy depends on the initial configuration of $\mathcal{S}_0$. In other words, with probability $1 - o(1)$ all vertices will play Strategy 0 or 1 after the first step. Thus, we need to determine the probability of each of them. We will argue as in Lemma 3.1.2; however, as we intend to use Lemma 3.2.1 we will express $n_i = (n-1)p_i$ whereby $p_i = \frac{1}{3} + \frac{c_i}{\sqrt{n}} + O(1/n)$. Using this, we write $n_0(v; i) = m_i + \delta_i$, where $m_i = \lceil n(v)p_i \rceil$ for $i = 0, 1$ and $m_2 = n(v) - m_0 - m_1$. With this notation, we write for $i = 0, 1$

$$T_0(v; i) = (m_0 + \delta_0)q_{i,0} + (m_1 + \delta_1)q_{i,1} + (m_2 + \delta_2)q_{i,2}.$$

But $\delta_2 = -\delta_0 - \delta_1$, whereby

$$T_0(v; i) = m_0 q_{i,0} + m_1 q_{i,1} + m_2 q_{i,2} + \delta_0(q_{i,0} - q_{i,2}) + \delta_1(q_{i,1} - q_{i,2}).$$

Now, using the definition of $m_i = n(v)p_i + r_i$, where $0 \le r_i < 1$, and the fact that $p_i = \frac{1}{3} + \frac{c_i}{\sqrt{n}} + O(1/n)$, we have that

$$m_0 q_{i,0} + m_1 q_{i,1} + m_2 q_{i,2} = \frac{n(v)}{3}\Sigma R_i + \frac{n(v)}{\sqrt{n}} \sum_{k=0}^{2} c_k q_{i,k} + \sum_{k=0}^{2} r_k q_{i,k} + O\left(\frac{n(v)}{n}\right).$$

Now, note that $c_0 + c_1 + c_2 = 0$ and also $r_2 = -r_0 - r_1$, the above becomes:

$$m_0 q_{i,0} + m_1 q_{i,1} + m_2 q_{i,2} = \frac{n(v)}{3} \Sigma R_i + \frac{n(v)}{\sqrt{n}} \left(c_0(q_{i,0} - q_{i,2}) + c_1(q_{i,1} - q_{i,2})\right) + O(1) + O\left(\frac{n(v)}{n}\right).$$

Therefore,

$$\begin{aligned} T_0(v;0) - T_0(v;1) &= \frac{n(v)}{3}(\Sigma R_0 - \Sigma R_1) + \frac{n(v)}{\sqrt{n}}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) \\ &\quad + \delta_0 C_{0,1}^{0,2} + \delta_1 C_{0,1}^{1,2} + O(1) + O\left(\frac{n(v)}{n}\right). \end{aligned} \tag{3.13}$$

But, by our hypothesis, $\Sigma R_0 - \Sigma R_1 = 0$ and therefore the payoff difference in (3.13) becomes

$$T_0(v;0) - T_0(v;1) = \delta_0 C_{0,1}^{0,2} + \delta_1 C_{0,1}^{1,2} + \frac{n(v)}{\sqrt{n}}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) + O(1) + O\left(\frac{n(v)}{n}\right). \tag{3.14}$$

So by (3.14), $T_0(v;0) - T_0(v,1) > 0$ if and only if

$$\delta_0 C_{0,1}^{0,2} + \delta_1 C_{0,1}^{1,2} > -\frac{n(v)}{\sqrt{n}}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) + O(1) + O\left(\frac{n(v)}{n}\right).$$

We re-scale both sides. Dividing by $n(v)^{1/2}$ and setting $x_i = \delta_i / n(v)^{1/2}$, the above is written as:

$$x_0 C_{0,1}^{0,2} + x_1 C_{0,1}^{1,2} > -\sqrt{\frac{n(v)}{n}}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) + O\left(\frac{n(v)^{1/2}}{n} + \frac{1}{n(v)^{1/2}}\right). \tag{3.15}$$

Recall that we have assumed that $|n(v) - np| \leq (np)^{1/2} \log n$. Thus, $\sqrt{n(v)/n} = p^{1/2}(1 + o(1))$ and $n(v)^{1/2}/n = O(p^{1/2}/n^{1/2}) = o(p^{1/2})$. Finally, since $1/n(v)^{1/2} = O(1/(np)^{1/2})$ and $1/n^{1/2} \ll p$, we also have that $1/n(v)^{1/2} = o(p^{1/2})$. Therefore, the above expression becomes:

$$x_0 C_{0,1}^{0,2} + x_1 C_{0,1}^{1,2} > -p^{1/2}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) + o(p^{1/2}).$$

We recall that $\mathbf{x} = (x_0, x_1)^T$ while we define $\mathbf{c} = (c_0, c_1)^T$ and $\mathbf{C} = \left(C_{0,1}^{0,2}, C_{0,1}^{1,2}\right)^T$, where the $c_i's$ are given by (3.2) and $C_{i,j}^{\ell,k}$ is defined as in Subsection 3.1.1. The above equation may also be rewritten as

$$\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}).$$

Therefore, we conclude that

$$\begin{aligned}
&\mathbb{P}\left[T_0(v;0) > T_0(v;1) \mid n(v), (S_0, S_1, S_2)\right] \\
&\quad = \mathbb{P}\left[x_0 C_{0,1}^{0,2} + x_1 C_{0,1}^{1,2} > -p^{1/2}\left(c_0 C_{0,1}^{0,2} + c_1 C_{0,1}^{1,2}\right) + o(p^{1/2}) \mid n(v), \mathcal{S}_0\right] \\
&\quad = \mathbb{P}\left[\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \mid n(v), \mathcal{S}_0\right]. \qquad (3.16)
\end{aligned}$$

Let $\mathcal{L}$ denote the set of points $\{n(v)^{-1/2} \cdot (k_0, k_1) : k_i \in \mathbb{Z}, i = 0, 1\}$ and let $\mathcal{L}'$ denote its restriction $\{(x_0, x_1) \in \mathcal{L} : |x_i| \leq \log n, i = 0, 1\}$. Using Lemma 3.2.1 and Corollary 3.1.4, we have the above probability is written as:

$$\begin{aligned}
&\mathbb{P}\left[\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \mid n(v), \mathcal{S}_0\right] \\
&\qquad = \frac{1}{n(v)}\left(1 + o(p^{1/2})\right) \sum_{\substack{\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \\ \mathbf{x} \in \mathcal{L}'}} \phi(\mathbf{x}) + O\left(e^{-(\log n)^2}\right).
\end{aligned}$$

Clearly $\exp(-\log n)^2 \ll p^{1/2}$ and we also observe by Lemma 3.2.9 and Lemma 3.4.2,

$$\left| o(p^{1/2}) \cdot \frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \\ \mathbf{x} \in \mathcal{L}'}} \phi(\mathbf{x}) \right| = o(p^{1/2}).$$

Therefore,

$$\mathbb{P}\left[\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \mid n(v), \mathcal{S}_0\right] = \frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) \\ \mathbf{x} \in \mathcal{L}'}} \phi(\mathbf{x}) + o(p^{1/2}).$$

We write

$$\frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}\in\mathcal{L}' \\ \mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2})}} \phi(\mathbf{x}) = \frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}\in\mathcal{L}' \\ \mathbf{x}^T\cdot\mathbf{C}\geq 0}} \phi(\mathbf{x}) + \frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}\in\mathcal{L}' \\ 0>\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2})}} \phi(\mathbf{x}).$$

We will show that for n sufficiently large:

$$\frac{1}{n(v)} \cdot \sum_{\mathbf{x}\in\mathcal{L}' \,:\, \mathbf{x}^T\cdot\mathbf{C}\geq 0} \phi(\mathbf{x}) = \frac{1}{2} + o(p^{1/2}) \tag{3.17}$$

and for some $\delta > 0$

$$\frac{1}{n(v)} \cdot \sum_{\substack{0>\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2}) \\ \mathbf{x}\in\mathcal{L}'}} \phi(\mathbf{x}) > \delta \cdot p^{1/2}. \tag{3.18}$$

So for n sufficiently large

$$\frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2}) \\ \mathbf{x}\in\mathcal{L}'}} \phi(\mathbf{x}) \geq \frac{1}{2} + \frac{\delta}{2} \cdot p^{1/2}. \tag{3.19}$$

While (3.19) suffices as a suitable lower bound, we also consider a more precise decomposition of this term. Namely,

$$\frac{1}{n(v)} \cdot \sum_{\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2}) \mathbf{x}\in\mathcal{L}'} \phi(\mathbf{x}) = \frac{1}{2} + \int_{\substack{0>\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2}) \\ \|\mathbf{x}\|_\infty<\log n}} \phi(\mathbf{x}) \mathrm{d}\mathbf{x} + o(p^{1/2}). \tag{3.20}$$

For n sufficiently large, equation (3.20) is a direct consequence of (3.17) and Lemma 3.2.8 given below. We refer to the integral in the above expression as $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$. We proceed with the proof of these inequalities. We lead by considering an approximation of the sum of $\phi(\mathbf{x})$ for points in $\mathcal{L}'$ to an appropriate Gaussian integral. For any $\mathbf{x} = (x_0, x_1)^T \in \mathcal{L}$, let

$\mathcal{B}_{\mathbf{x}}$ denote the box $[x_0, x_0 + n(v)^{-1/2}) \times [x_1, x_1 + n(v)^{-1/2})$.

Lemma 3.2.8. *Suppose* $\mathbf{x} \in \mathcal{L}$. *Then for any* $\mathbf{x}', \mathbf{x}'' \in \mathcal{B}_{\mathbf{x}}$ *we have uniformly over any choice of* $\mathbf{x}, \mathbf{x}'$ *and* $\mathbf{x}''$,

$$|\phi(\mathbf{x}') - \phi(\mathbf{x}'')| = O\left(\frac{1}{\sqrt{n(v)}}\right).$$

Proof. Suppose we have $\mathbf{x}' = (x_0', x_1')^T$ and $\mathbf{x}'' = (x_0'', x_1'')^T$, both belonging to $\mathcal{B}_{\mathbf{x}}$. We appeal to the multivariate mean value theorem as follows: there exists some $c \in (0,1)$ such that,

$$|\phi(\mathbf{x}') - \phi(\mathbf{x}'')| \leq \|\nabla\phi((1-c)\mathbf{x}' + c\mathbf{x}'')\|_2 \|\mathbf{x}' - \mathbf{x}''\|_2. \tag{3.21}$$

We remark that equation (3.21) is a direct consequence of applying the usual single variable mean value theorem to ϕ, over the line segment with endpoints $\mathbf{x}'$ and $\mathbf{x}''$. We set $z_0 = x_0'(1-c) + x_0''c$ and $z_1 = x_1'(1-c) + x_1''c$, and consider the order of $\|\nabla\phi((z_0, z_1))\|_2$. We remark that the exponent in $\phi(z_0, z_1)$ is a quadratic form in $\mathbf{z} = (z_0, z_1)^T$ thus we have that,

$$-\frac{1}{2}\mathbf{z}^T\Sigma\mathbf{z} = h(z_0, z_1),$$

where h quadratic in z_0 and z_1. Hence it follows from the definition of ϕ that,

$$\|\nabla\phi((z_0, z_1))\|_2 = |\phi(\mathbf{z})| \|\nabla h((z_0, z_1)\|_2.$$

Now we observe that as h is quadratic in z_0 and z_1 it follows that,

$$\|\nabla h((z_0, z_1)\|_2 = O(\max\{|z_0|^2, |z_1|^2\}) = O(1).$$

We also have that $\phi(\mathbf{z}) \leq 1$. Hence we have that,

$$|\phi(\mathbf{x}') - \phi(\mathbf{x}'')| = O(\|\mathbf{x}' - \mathbf{x}''\|_2).$$

However, as $\mathbf{x}'$ and $\mathbf{x}'' \in \mathcal{B}_{\mathbf{x}}$ then we have that $\|\mathbf{x}' - \mathbf{x}''\|_2 = O\left(1/\sqrt{n(v)}\right)$, which concludes the proof of this Lemma. □

By applying the above, we can consider an approximation of our sums by a Gaussian integral. We define $\mathcal{B}^{\infty}_{\log n} = \{\mathbf{z} = (z_0, z_1) \in \mathbb{R}^2 : \|\mathbf{z}\|_\infty < \log n\}$.

Lemma 3.2.9. *Suppose $\mathcal{D} \subset \mathcal{B}^{\infty}_{\log n}$ is a Borel set. Then we have*

$$\left| \frac{1}{n(v)} \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{x}) - \int_{\mathcal{D}} \phi(\mathbf{z}) \mathrm{d}\mathbf{z} \right| = o(p^{1/2}).$$

Proof. We deduce an asymptotic for the above in terms of an appropriate Gaussian integral. For each $\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'$ we define

$$\mathbf{z}_{\max}(\mathbf{x}) = \operatorname*{argmax}_{\mathbf{z} \in \mathcal{B}_{\mathbf{x}}} \{\phi(\mathbf{z})\} \ \text{ and, } \ \mathbf{z}_{\min}(\mathbf{x}) = \operatorname*{argmin}_{\mathbf{z} \in \mathcal{B}_{\mathbf{x}}} \{\phi(\mathbf{z})\},$$

the locations of the maximum and minimum values of ϕ across the box region $\mathcal{B}_{\mathbf{x}}$. Thus we will now deduce upper and lower bounds in terms of the Gaussian integral over $\mathcal{D}$. By Lemma 3.2.8 we have,

$$\begin{aligned} \frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{x}) &= \frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{z}_{\min}(\mathbf{x})) + \frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} |\phi(\mathbf{x}) - \phi(\mathbf{z}_{\min}(\mathbf{x}))| \\ &= \frac{1}{n(v)} \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{z}_{\min}(\mathbf{x})) + O\left(\frac{1}{n(v)^{3/2}} |\mathcal{D} \cap \mathcal{L}'| \right). \end{aligned}$$

We observe that $|\mathcal{D} \cap \mathcal{L}'| \leq |\mathcal{L}'| = O(n(v) \log^2 n)$. Therefore we observe that as $p \gg n^{-1/2} \log^3 n$ it follows that $np \gg p$ and, hence

$$\frac{|\mathcal{D} \cap \mathcal{L}'|}{n(v)^{3/2}} = O\left(\frac{\log n}{\sqrt{n(v)}} \right) = o(p^{1/2}).$$

Thus by taking an appropriate bound using a Riemann integral across $\mathcal{B}_{\mathbf{x}}$ we have,

$$\frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{x}) \leq \int_{\mathcal{D}} \phi(\mathbf{x}) d\mathbf{x} + o(p^{1/2}).$$

Similarly, for a lower bound, we have,

$$\begin{aligned}\frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{x}) &= \frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{z}_{\max}(\mathbf{x})) - \frac{1}{n(v)} \cdot \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} |\phi(\mathbf{z}_{\max}(\mathbf{x})) - \phi(\mathbf{x})| \\ &= \frac{1}{n(v)} \sum_{\mathbf{x} \in \mathcal{D} \cap \mathcal{L}'} \phi(\mathbf{z}_{\max}(\mathbf{x})) - O\left(\frac{1}{n(v)^{3/2}} |\mathcal{D} \cap \mathcal{L}'|\right) \\ &\geq \int_{\mathcal{D}} \phi(\mathbf{x}) d\mathbf{x} - o(p^{1/2}). \qquad \square\end{aligned}$$

Proof of equation (3.17)

For $\mathbf{C}$ as given above, we define the sets,

$$\mathcal{C}^+ = \{\mathbf{z} = (z_0, z_1) \in \mathcal{B}^{\infty}_{\log n} : \mathbf{z} \cdot \mathbf{C} \geq 0\} \text{ and } \mathcal{C}^- = \{\mathbf{z} = (z_0, z_1) \in \mathcal{B}^{\infty}_{\log n} : \mathbf{z} \cdot \mathbf{C} < 0\}.$$

By Lemma 3.2.9,

$$\left| \frac{1}{n(v)} \cdot \sum_{\substack{\mathbf{x}^T \cdot \mathbf{C} > 0 \\ \mathbf{x} \in \mathcal{L}'}} \phi(\mathbf{x}) - \int_{\mathcal{C}^+} \phi(\mathbf{z}) \mathrm{d}\mathbf{z} \right| = o(p^{1/2}).$$

Thus, it suffices for us to show that the above integral is equal to $1/2 + o(p^{1/2})$. Firstly we observe that as Σ is symmetric and semi-positive definite it follows that

$$\int_{\mathbb{R}^2} \phi(\mathbf{z}) \mathrm{d}\mathbf{z} = 1.$$

Furthermore we have by Lemma 3.4.2,

$$\int_{\mathbf{z} \notin \mathcal{B}^{\infty}_{\log n}} \phi(\mathbf{z}) \mathrm{d}\mathbf{z} = O\left(e^{-(\log n)^2}\right) = o(p^{1/2}).$$

However we also observe that $\phi(\mathbf{z}) = \phi(-\mathbf{z})$, hence combining the above we obtain the following two equations,

$$\int_{\mathcal{C}^+} \phi(\mathbf{z}) \mathrm{d}\mathbf{z} = \int_{\mathcal{C}^-} \phi(\mathbf{z}) \mathrm{d}\mathbf{z},$$

$$\int_{\mathcal{C}^+} \phi(\mathbf{z})\mathrm{d}\mathbf{z} + \int_{\mathcal{C}^-} \phi(\mathbf{z})\mathrm{d}\mathbf{z} = 1 - O\left(e^{-(\log n)^2}\right)$$

Combining the above gives the desired conclusion.

Proof of equation (3.18)

Let $\mathcal{L}'' = \{(x_0, x_1) \in \mathcal{L}' \ : \ |x_0| \leq 1, \ |x_1| \leq |C_0/C_1| + 2\|\mathbf{C}\|_\infty/\varepsilon\}$. As $\mathcal{L}'' \subset \mathcal{L}'$, provided that n is sufficiently large, we first bound

$$\frac{1}{n(v)} \cdot \sum_{\substack{0 > \mathbf{x}^T\cdot\mathbf{C} > -p^{1/2}\mathbf{c}^T\cdot\mathbf{C} + o(p^{1/2}) \\ \mathbf{x}\in\mathcal{L}'}} \phi(\mathbf{x}) > \frac{1}{n(v)} \cdot \sum_{\substack{0 > \mathbf{x}^T\cdot\mathbf{C} > -p^{1/2}\mathbf{c}^T\cdot\mathbf{C} + o(p^{1/2}) \\ \mathbf{x}\in\mathcal{L}''}} \phi(\mathbf{x}).$$

Let $\gamma > 0$ be such that $\phi(\mathbf{x}) > \gamma$, for any $\mathbf{x} \in \mathcal{L}''$. Thus, we can further bound:

$$\frac{1}{n(v)} \cdot \sum_{\substack{0 > \mathbf{x}^T\cdot\mathbf{C} > -p^{1/2}\mathbf{c}^T\cdot\mathbf{C} + o(p^{1/2}) \\ \mathbf{x}\in\mathcal{L}''}} \phi(\mathbf{x}) > \frac{\gamma}{n(v)} \cdot |\{\mathbf{x} \in \mathcal{L}'' \ : \ 0 > \mathbf{x}^T\cdot\mathbf{C} > -p^{1/2}\mathbf{c}^T\cdot\mathbf{C} + o(p^{1/2})\}|. \tag{3.22}$$

We will now bound from below the size of the set $\{\mathbf{x} \in \mathcal{L}'' \ : \ 0 > \mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2})\}$. Set $\alpha = p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2}) > 0$. Since $\mathbf{c} = (c_0, c_1)^T$ with $|c_0|, |c_1| \leq 1/\varepsilon$ and $p < 1$, for any n sufficiently large we have

$$(1 + o(1))p^{1/2}\varepsilon\|\mathbf{C}\|_\infty < |\alpha| < |\mathbf{c}^T \cdot \mathbf{C}| \leq \|\mathbf{c}\|_2\|\mathbf{C}\|_2 \leq \frac{2}{\varepsilon}\|\mathbf{C}\|_\infty.$$

With this notation, note that the above set consists of points $n(v)^{-1/2} \cdot (i, j)$, with $i, j \in \mathbb{Z}$, that satisfy

$$0 > \frac{1}{n(v)^{1/2}}\left(iC_0 + jC_1\right) > -\alpha.$$

Thus, for any $i \in \mathbb{Z}$ with $|i| \leq n(v)^{1/2}$, there are at least

$$\frac{1}{2}n(v)^{1/2}\frac{|\alpha|}{|C_1|} > \frac{1 + o(1)}{2}n(v)^{1/2}p^{1/2}\varepsilon\|\mathbf{C}\|_\infty$$

choices for j such that $n(v)^{-1/2}(i,j) \in \mathcal{L}''$. Therefore,

$$|\{\mathbf{x} \in \mathcal{L}'' \ : \ 0 > \mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2})\}| \geq \frac{1}{8}p^{1/2}n(v)\varepsilon\|\mathbf{C}\|_\infty.$$

So the left-hand side of (3.22) is bounded from below by $\frac{\gamma\varepsilon\|\mathbf{C}\|_\infty}{8} \cdot p^{1/2}$, thus concluding the proof of (3.18). We observe Lemma 3.2.8 implies that,

$$\beta_{\mathbf{c}^T \cdot \mathbf{C}} > \delta \cdot p^{1/2} + o(\sqrt{p}). \tag{3.23}$$

Concentration after the first step

Let X_0 denote the random variable that is the number of vertices that adopt Strategy 0, given that $\mathbf{c}^T \cdot \mathbf{C} > 0$. By Lemma 3.4.1 (found in the auxiliary results section), we may assume this as $\mathbf{c}^T \cdot \mathbf{C} \neq 0$ with high probability. By the previous analysis, there exists δ such that for any n sufficiently large

$$\mathbb{E}\left(X_0\right) = \frac{n}{2} + n\beta_{\mathbf{c}^T \cdot \mathbf{C}} + o(np^{1/2}) \geq \frac{n}{2} + \delta np^{1/2}. \tag{3.24}$$

We will show that with probability $1 - o(1)$ for n sufficiently large we have

$$\frac{n}{2} + n\beta_{\mathbf{c}^T \cdot \mathbf{C}} - \frac{\delta}{3}np^{1/2} \leq X_0 \leq \frac{n}{2} + n\beta_{\mathbf{c} \cdot \mathbf{C}} + \frac{\delta}{3}np^{1/2}. \tag{3.25}$$

and consequently by (3.23),

$$X_0 \geq \frac{n}{2} + \frac{\delta}{2}np^{1/2}. \tag{3.26}$$

To this end, we will show that

$$\mathbb{P}\left[|X_0 - \mathbb{E}\left(X_0\right)| > \frac{\delta}{4}np^{1/2}\right] = o(1), \tag{3.27}$$

We will use a second-moment argument - Chebyschev's inequality yields

$$\mathbb{P}\left[|X_0 - \mathbb{E}\left(X_0\right)| > \frac{\delta}{4}np^{1/2}\right] \leq \frac{16}{\delta^2} \cdot \frac{\mathrm{Var}(X_0)}{n^2 p}.$$

We will bound $\mathrm{Var}(X_0)$ and, in particular, we will show that

$$\mathrm{Var}(X_0) = O\left(n^{3/2}\right). \tag{3.28}$$

This implies that

$$\mathbb{P}\left[|X_0 - \mathbb{E}\left(X_0\right)| > \frac{\delta}{4}np^{1/2}\right] = O\left(\frac{1}{n^{1/2}p}\right) = o(1).$$

Hence we have with high probability,

$$|X_0 - \mathbb{E}\left(X_0\right)| < \frac{\delta}{4}np^{1/2}$$

and thus,

$$\frac{n}{2} + n\beta_{\mathbf{c}^T \cdot \mathbf{C}} - \frac{\delta}{4}np^{1/2} + o(n\sqrt{p}) \leq X_0 \leq \frac{n}{2} + n\beta_{\mathbf{c}^T \cdot \mathbf{C}} + \frac{\delta}{4}np^{1/2} + o(n\sqrt{p}). \tag{3.29}$$

Taking n sufficiently large implies the bounds given by (3.25).

Proof of (3.28). We write

$$\mathbb{E}\left(X_0\right) = \sum_{v \in V_n} \mathbb{P}\left[S_1(v) = 0\right].$$

and

$$\mathrm{Var}(X_0) \leq \sum_{v \neq v'} \left(\mathbb{P}\left[S_1(v) = S_1(v') = 0\right] - \mathbb{P}\left[S_1(v) = 0\right]\mathbb{P}\left[S_1(v') = 0\right]\right) + \mathbb{E}\left(X_0\right).$$

We will show that uniformly over all distinct $v, v' \in V_n$ we have

$$\mathbb{P}\left[S_1(v) = S_1(v') = 0\right] - \mathbb{P}\left[S_1(v) = 0\right]\mathbb{P}\left[S_1(v') = 0\right] = O\left(\frac{1}{n^{1/2}}\right). \tag{3.30}$$

Using the trivial bound that $\mathbb{E}(X_0) \leq n$ we obtain:

$$\mathrm{Var}(X_0) = O\left(\frac{n^2}{n^{1/2}}\right) + n = O\left(n^{3/2}\right).$$

□

For any $v \in V_n$ let J_v denote the indicator random variable that is equal to 1 precisely on $\{S_1(v) = 0\}$, thus we note that $J_v = 1_{S_1(v)=0}$. We will bound

$$\mathbb{E}\left(J_v \cdot J_{v'}\right) - \mathbb{E}\left(J_v\right) \cdot \mathbb{E}\left(J_{v'}\right).$$

We define $\hat{n}(v) = |N_G(v) \setminus \{v, v'\}|$. Applying the law of total probability and conditioning on whether or not $v \sim v'$, as well as on the values of $\hat{n}(v)$ and $\hat{n}(v')$, it follows that:

$$\mathbb{E}\left(J_v \cdot J_{v'}\right) = \\ p \cdot \mathbb{E}\left(\mathbb{E}\left(J_v \cdot J_{v'} \mid \hat{n}(v), \hat{n}(v')\right) \mid v \sim v'\right) + (1-p) \cdot \mathbb{E}\left(\mathbb{E}\left(J_v \cdot J_{v'} \mid \hat{n}(v), \hat{n}(v')\right) \mid v \not\sim v'\right). \tag{3.31}$$

To begin, we will start with the first summand on the right-hand side. Recall that we are conditioning on the initial configuration $\mathcal{S}_0 = \left(S_0^{(0)}, S_0^{(1)}, S_0^{(2)}\right)$ being such that $|S_0^{(j)}| = n/3 + c_j\sqrt{n}$, for $j \in \{0, 1, 2\}$, where $\varepsilon < |c_j| < 1/\varepsilon$.

Having selected v and v', let us abbreviate $\mathbb{E}_\sim[\cdot] = \mathbb{E}\left(\cdot \mid v \sim v'\right)$. Using this shorthand, we can write explicitly

$$\mathbb{E}_\sim\left[\mathbb{E}\left(J_v \cdot J_{v'} \mid \hat{n}(v), \hat{n}(v')\right)\right] =$$

$$\sum_{\hat{k}=0}^{n-2}\sum_{\hat{k}'=0}^{n-2}\mathbb{E}_{\sim}\left[J_v \cdot J_{v'} \mid \hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right].$$

Note that conditional on $v \sim v'$ and on $\hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'$, the random variables J_v and $J_{v'}$ are independent. The random variables $\hat{n}(v), \hat{n}(v')$ are independent too. Furthermore, J_v depends only $N(v)\setminus\{v,v'\}$ and $J_{v'}$ depends only on $N(v')\setminus\{v,v'\}$, if we condition on $v \sim v'$. So each summand in the above double sum can be written as:

$$\mathbb{E}_{\sim}\left[J_v \cdot J_{v'} \mid \hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right]=$$
$$\mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{E}_{\sim}\left[J_{v'} \mid \hat{n}(v)=\hat{k}, \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]\mathbb{P}\left[\hat{n}(v')=\hat{k}'\right]=$$
$$\mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}\right]\cdot\mathbb{E}_{\sim}\left[J_{v'} \mid \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]\mathbb{P}\left[\hat{n}(v')=\hat{k}'\right].$$

Therefore, the double sum itself can be factorised as follows:

$$\mathbb{E}_{\sim}\left[\mathbb{E}\left(J_v \cdot J_{v'} \mid \hat{n}(v), \hat{n}(v')\right)\right]=$$
$$=\left(\sum_{\hat{k}=0}^{n-2}\mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]\right)\cdot\left(\sum_{\hat{k}'=0}^{n-2}\mathbb{E}_{\sim}\left[J_{v'} \mid \hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v')=\hat{k}'\right]\right). \tag{3.32}$$

Let us fix $\hat{k}\in\{0,\dots,n-2\}$ and define

$$\hat{\mathcal{K}}_{\hat{k}}=\left\{(\hat{k}_0,\hat{k}_1,\hat{k}_2)\in\mathbb{N}_0^3 \ : \ \sum_{j=0}^{2}\hat{k}_j=\hat{k}, \ \sum_{j=0}^{2}\hat{k}_j(q_{0j}-q_{1j})>-(q_{0,S(v')}-q_{1,S(v')})\right\}.$$

Using this, we express:

$$\mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}\right]=\sum_{(\hat{k}_0,\hat{k}_1,\hat{k}_2)\in\hat{\mathcal{K}}_{\hat{k}}}\mathbb{P}\left[\hat{n}_0(v;j)=\hat{k}_j, \ j=0,1,2 \mid \hat{n}(v)=\hat{k}\right]. \tag{3.33}$$

Now, set $k_j=\hat{k}_j+1_{S(v')=j}$, for $j=0,1,2$. We will compare $\mathbb{P}\left[\hat{n}_0(v;j)=\hat{k}_j, \ j=0,1,2\right]$ with $\mathbb{P}\left[n_0(v;j)=k_j, \ j=0,1,2\right]$. For any $v\in V_n$ let $\mathcal{K}_k=\{(k_0,k_1,k_2)\in\mathbb{N}_0^3 \ : \ \sum_{j=0}^{2}k_j=$

$k+1,\ \sum_{j=0}^{2} k_j(q_{0j}-q_{1j})>0\}$.

Claim 3.2.10. *For any $k \in \{0,\dots,n-1\}$ and all $(k_0,k_1,k_2) \in \mathcal{K}_k$ we have*

$$\frac{\binom{n_0-1_{s(v)=0}-1_{s(v')=0}}{k_0-1_{s(v')=0}}\binom{n_1-1_{s(v)=1}-1_{s(v')=1}}{k_1-1_{s(v')=1}}\binom{n_2-1_{s(v)=2}-1_{s(v')=2}}{k_2-1_{s(v')=2}}}{\binom{n-2}{k}} \cdot \left(\frac{\binom{n_0-1_{s(v)=0}}{k_0}\binom{n_1-1_{s(v)=1}}{k_1}\binom{n_2-1_{s(v)=2}}{k_2}}{\binom{n-1}{k+1}}\right)^{-1}$$
$$= \frac{n-1}{k+1}\cdot\frac{k_{s(v')}}{n_{s(v')}-1_{s(v')=s(v)}}.$$

Proof. Note that if $s(v')=j$

$$\binom{n_j-1_{s(v)=j}-1_{s(v')=j}}{k_j-1_{s(v')=j}}\cdot\binom{n_j-1_{s(v)=j}}{k_j}^{-1} = \frac{k_j}{n_j-1_{s(v)=j}}$$

but otherwise

$$\binom{n_j-1_{s(v)=j}-1_{s(v')=j}}{k_j-1_{s(v')=j}}\cdot\binom{n_j-1_{s(v)=j}}{k_j}^{-1} = 1.$$

Similarly,

$$\binom{n-1}{k+1}\cdot\binom{n-2}{k}^{-1} = \frac{n-1}{k+1}.$$

So the claim follows. □

Therefore,

$$\mathbb{P}\left[\hat{n}_0(v;j)=\hat{k}_j,\ j=0,1,2 \mid \hat{n}(v)=\hat{k}\right] =$$
$$\frac{n-1}{\hat{k}+1}\cdot\frac{k_{s(v')}}{n_{s(v')}-1_{s(v')=s(v)}}\cdot\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}+1\right].$$

Now, let us relate $\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]$ with $\mathbb{P}\left[n(v)=\hat{k}+1\right]$:

$$\begin{aligned}
\mathbb{P}\left[\hat{n}(v)=\hat{k}\right] &= \binom{n-2}{\hat{k}}p^{\hat{k}}(1-p)^{n-2-\hat{k}} \\
&= \frac{1}{p}\frac{\hat{k}+1}{n-1}\cdot\binom{n-1}{\hat{k}+1}p^{\hat{k}+1}.(1-p)^{n-1-(\hat{k}+1)} \\
&= \frac{1}{p}\frac{\hat{k}+1}{n-1}\cdot\mathbb{P}\left[n(v)=\hat{k}+1\right].
\end{aligned}$$

With these, we can write

$$\sum_{\hat{k}=0}^{n-2} \mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}\right]\cdot \mathbb{P}\left[\hat{n}(v)=\hat{k}\right] =$$
$$\sum_{\hat{k}=0}^{n-2} \mathbb{P}\left[n(v)=\hat{k}+1\right] \sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}+1}} \frac{k_{s(v')}}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}+1\right]$$

To this end, we express $k_{s(v')} = (n_{s(v')}-1_{s(v')=s(v)})p+\delta_{s(v')}$. Using this, we get:

$$\sum_{\hat{k}=0}^{n-2} \mathbb{E}_{\sim}\left[J_v \mid \hat{n}(v)=\hat{k}\right]\cdot \mathbb{P}\left[\hat{n}(v)=\hat{k}\right] =$$
$$\sum_{\hat{k}=0}^{n-2} \sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}+1}} \mathbb{P}\left[n(v)=\hat{k}+1\right]\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}+1\right]+$$
$$\sum_{\hat{k}=0}^{n-2} \mathbb{P}\left[n(v)=\hat{k}+1\right] \sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}+1}} \frac{\delta_{s(v')}}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}+1\right].$$

The first term is precisely equal to $\mathbb{E}\left(J_v\right)$, while, for the second term, we combine:

$$\mathbb{P}\left[n(v)=\hat{k}+1\right]\cdot\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}+1\right] = \mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2\right].$$

Note that the random variables $n_0(v;j)$ for $j=0,1,2$ are independent and binomially distributed. Let $\Gamma_k = \{(k_0,k_1,k_2)\in\mathbb{N}_0^3 : \sum_{j=0}^{2} k_j = k, 0\le k_j \le n-1\}$. Using this notation and the triangle inequallity we bound,

$$\left|\sum_{\hat{k}=0}^{n-2} \sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}+1}} \frac{\delta_{s(v')}}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2\right]\right| \le$$
$$\le \sum_{\hat{k}=0}^{n-2} \sum_{(k_0,k_1,k_2)\in\Gamma_{\hat{k}+1}} \frac{|\delta_{s(v')}|}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2\right]$$
$$\le \sum_{k=0}^{n-1} \sum_{(k_0,k_1,k_2)\in\Gamma_{k}} \frac{|\delta_{s(v')}|}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2\right]$$

As we work in Γ_k we may treat this sum as three independent binomials hence we have the above is equal to,

$$\sum_{k_0=0}^{n_0}\sum_{k_1=0}^{n_1}\sum_{k_2=0}^{n_2}\frac{|\delta_{s(v')}|}{(n_{s(v')}-1_{s(v')=s(v)})p}\prod_{j=0}^{2}\mathbb{P}\left[n_0(v;j)=k_j\right]$$
$$=\sum_{k_{s(v')}=0}^{n_{s(v')}}\frac{|\delta_{s(v')}|}{(n_{s(v')}-1_{s(v')=s(v)})p}\mathbb{P}\left[n_0(v;s(v'))=k_{s(v')}\right]\overset{n_{s(v')}=np_{s(v')}}{=}O\left(\frac{1}{\sqrt{np}}\right).$$

Therefore,

$$\sum_{\hat{k}=0}^{n-2}\mathbb{E}_{\sim}\left[J_v\mid\hat{n}(v)=\hat{k}\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]\le\mathbb{E}\left(J_v\right)+O\left(\frac{1}{\sqrt{np}}\right).$$

Using this upper bound in (3.32), we get

$$\mathbb{E}_{\sim}\left[\mathbb{E}\left(J_v\cdot J_{v'}\mid\hat{n}(v),\hat{n}(v')\right)\right]\le\mathbb{E}\left(J_v\right)\cdot\mathbb{E}\left(J_{v'}\right)+O\left(\frac{1}{\sqrt{np}}\right).\tag{3.34}$$

Now, let us abbreviate $\mathbb{E}_{\not\sim}[\cdot]=\mathbb{E}\left(\cdot\mid v\not\sim v'\right)$. Using the law of total probability conditioning again on the values of $\hat{n}(v)$ and $\hat{n}(v')$, we get

$$\mathbb{E}_{\not\sim}\left[\mathbb{E}\left(J_v\cdot J_{v'}\mid\hat{n}(v),\hat{n}(v')\right)\right]=$$
$$\sum_{\hat{k}=0}^{n-2}\sum_{\hat{k}'=0}^{n-2}\mathbb{E}_{\not\sim}\left[J_v\cdot J_{v'}\mid\hat{n}(v)=\hat{k},\hat{n}(v')=\hat{k}'\right]\cdot\mathbb{P}\left[\hat{n}(v)=\hat{k},\hat{n}(v')=\hat{k}'\right].$$

We now show a similar relation for $\mathbb{E}_{\not\sim}\left[J_v\cdot J_{v'}\mid\hat{n}(v)=\hat{k},\hat{n}(v')=\hat{k}'\right]$. Note that

$$\mathbb{E}_{\not\sim}\left[J_v\cdot J_{v'}\mid\hat{n}(v)=\hat{k},\hat{n}(v')=\hat{k}'\right]=\mathbb{E}_{\not\sim}\left[J_v\mid\hat{n}(v)=\hat{k}\right]\cdot\mathbb{E}_{\not\sim}\left[J_{v'}\mid\hat{n}(v')=\hat{k}'\right].$$

We deal with $\mathbb{E}_{\not\sim}\left[J_v\mid\hat{n}(v)=\hat{k}\right]$, as the analogous calculation can be carried out for v'. Note that Setting $\mathcal{K}_{\hat{k}}=\{(k_0,k_1,k_2)\ :\ \sum_{j=0}^{2}k_j=\hat{k},\ \sum_{j=0}^{2}k_j(q_{0j}-q_{1j})>0\}$ can write

$$\mathbb{E}_{\not\sim}\left[J_v\mid\hat{n}(v)=\hat{k}\right]=\sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}}}\mathbb{P}\left[\hat{n}_0(v;j)=k_j,\ j=0,1,2\mid\hat{n}(v)=\hat{k}\right].\tag{3.35}$$

We now show the analogue of Claim 3.2.10. In particular, we will relate the value of $\mathbb{P}\left[\hat{n}_0(v;j) = k_j,\ j = 0,1,2 \mid \hat{n}(v) = \hat{k}\right]$ to $\mathbb{P}\left[n_0(v;j) = k_j,\ j = 0,1,2 \mid n(v) = \hat{k}\right]$.

Claim 3.2.11. *Uniformly over all $(k_0, k_1, k_2) \in \mathcal{K}_{\hat{k}}$ we have*

$$\frac{\binom{n_0 - 1_{s(v)=0} - 1_{s(v')=0}}{k_0}\binom{n_1 - 1_{s(v)=1} - 1_{s(v')=1}}{k_1}\binom{n_2 - 1_{s(v)=2} - 1_{s(v')=2}}{k_2}}{\binom{n-2}{\hat{k}}} \cdot \left(\frac{\binom{n_0 - 1_{s(v)=0}}{k_0}\binom{n_1 - 1_{s(v)=1}}{k_1}\binom{n_2 - 1_{s(v)=2}}{k_2}}{\binom{n-1}{\hat{k}}}\right)^{-1}$$
$$= \frac{n-1}{n-\hat{k}-1} \cdot \frac{n_{s(v')} - k_{s(v')} - 1_{s(v')=s(v)}}{n_{s(v')} - 1_{s(v')=s(v)}}.$$

Proof. Again we follow a similar approach to that in Claim 3.2.10. We split into two cases. Firstly suppose $s(v) = s(v') = j$ for some $j \in \{0, 1, 2\}$. Then we have:

$$\binom{n_j - 1_{s(v)=j} - 1_{s(v')=j}}{k_j} \cdot \binom{n_j - 1_{s(v)=j}}{k_j}^{-1} = \frac{n_j - k_j - 1}{n_j - 1}.$$

While for any $i \neq j$ we have that:

$$\binom{n_i - 1_{s(v)=i} - 1_{s(v')=i}}{k_i} \cdot \binom{n_i - 1_{s(v)=i}}{k_i}^{-1} = 1.$$

Furthermore, we have that:

$$\binom{n-1}{\hat{k}} \cdot \binom{n-2}{\hat{k}-1}^{-1} = \frac{n-1}{\hat{k}},$$

and hence the result follows. In the case that $s(v) = i$ and $s(v') = j$ with $i \neq j$ then we have the following:

$$\binom{n_i - 1_{s(v)=i} - 1_{s(v')=i}}{k_i} \cdot \binom{n_i - 1_{s(v)=i}}{k_i}^{-1} = 1,$$

while,

$$\binom{n_j - 1_{s(v)=j} - 1_{s(v')=j}}{k_j} \cdot \binom{n_j - 1_{s(v)=j}}{k_j}^{-1} = \binom{n_j - 1}{k_j} \cdot \binom{n_j}{k_j}^{-1} = \frac{n_j - k_j}{n_j}.$$

Hence the result follows as $j = s(v')$.

$\square$

Therefore,

$$\mathbb{P}\left[\hat{n}_0(v;j)=\hat{k}_j,\ j=0,1,2 \mid \hat{n}(v)=\hat{k}\right]=$$
$$\frac{n-1}{n-\hat{k}-1}\cdot\frac{n_{s(v')}-k_{s(v')}-1_{s(v')=s(v)}}{n_{s(v')}-1_{s(v')=s(v)}}\cdot\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right]. \tag{3.36}$$

Furthermore,

$$\begin{aligned}\mathbb{P}\left[\hat{n}(v)=\hat{k}\right] &= \binom{n-2}{\hat{k}}p^{\hat{k}}(1-p)^{n-2-\hat{k}}\\ &= \frac{1}{1-p}\frac{n-\hat{k}-1}{n-1}\cdot\binom{n-1}{\hat{k}}p^{\hat{k}}.(1-p)^{n-1-\hat{k}}\\ &= \frac{1}{1-p}\frac{n-\hat{k}-1}{n-1}\cdot\mathbb{P}\left[n(v)=\hat{k}\right].\end{aligned}$$

Therefore,

$$\mathbb{P}\left[\hat{n}(v)=\hat{k}\right]\mathbb{P}\left[\hat{n}_0(v;j)=\hat{k}_j,\ j=0,1,2 \mid \hat{n}(v)=\hat{k}\right]=$$
$$=\frac{n_{s(v')}-k_{s(v')}-1_{s(v')=s(v)}}{(n_{s(v')}-1_{s(v')=s(v)})(1-p)}\cdot\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right]\mathbb{P}\left[n(v)=\hat{k}\right].$$

Now, we express $k_{s(v')}=p\left(n_{s(v')}-1_{s(v)=s(v')}\right)+\delta_{s(v')}$, where $n_{s(v')}=np_{s(v')}$. We observe,

$$\begin{aligned}&\sum_{\hat{k}=0}^{n-2}\mathbb{P}\left[n(v)=\hat{k}\right]\times\\ &\qquad\sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}}}\frac{n_{s(v')}-k_{s(v')}-1_{s(v')=s(v)}}{(n_{s(v')}-1_{s(v')=s(v)})(1-p)}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right]\\ =&\sum_{\hat{k}=0}^{n-2}\mathbb{P}\left[n(v)=\hat{k}\right]\times\\ &\qquad\sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}}}\frac{(n_{s(v')}-1_{s(v')=s(v)})(1-p)+\delta_{s(v')}}{(n_{s(v')}-1_{s(v')=s(v)})(1-p)}\mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right].\end{aligned}$$

$$(3.37)$$

As in the above calculation, the first summand is,

$$\sum_{\hat{k}=0}^{n-2} \mathbb{P}\left[n(v)=\hat{k}\right] \sum_{(k_0,k_1,k_2)\in\mathcal{K}_{\hat{k}}} \mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right] \leq \mathbb{E}\left(J_v\right). \quad (3.38)$$

By an analogous calculation to the above, the second summand is bounded above by,

$$\sum_{(k_0,k_1,k_2)\in\Gamma_k} \frac{|\delta_{s(v')}|}{(n_{s(v')}-1_{s(v')=s(v)})(1-p)} \mathbb{P}\left[n_0(v;j)=k_j,\ j=0,1,2 \mid n(v)=\hat{k}\right] = O\left(\sqrt{\frac{p}{n}}\right).$$

This implies that

$$\begin{aligned} \mathbb{E}\left(\mathbb{E}_{\not\sim}\left[J_v \cdot J_{v'} \mid \hat{n}(v), \hat{n}(v')\right]\right) &= \mathbb{E}\left(\mathbb{E}_{\not\sim}\left[J_v \mid \hat{n}(v)\right]\right) \cdot \mathbb{E}\left(\mathbb{E}_{\not\sim}\left[J_{v'} \mid \hat{n}(v')\right]\right) \\ &\leq \mathbb{E}\left(J_v\right) \cdot \mathbb{E}\left(J_{v'}\right) + O\left(\sqrt{\frac{p}{n}}\right). \end{aligned} \quad (3.39)$$

Using the bounds of (3.34) and (3.39) into (3.31) we deduce that

$$\mathbb{E}\left(J_v \cdot J_{v'}\right) \leq \mathbb{E}\left(J_v\right) \cdot \mathbb{E}\left(J_{v'}\right) + O\left(\frac{1}{n^{1/2}}\right).$$

3.2.4 After the First Round

In the previous subsection, we used a second-moment argument in order to show that in the case $\mathbf{c}^T \cdot \mathbf{C} > 0$, we have with probability $1-o(1)$ that $|S_1(0)| = \frac{n}{2} + a(n)$, for $a(n) \geq \frac{\delta}{2} np^{1/2}$. We recall that $|S_1(1)|$ denotes the number of vertices which adopt strategy 1 after the execution of the first round. On the event that $a(n) \geq \frac{\delta}{2} np^{1/2}$ we have, $|S_1(1)| \leq \frac{n}{2} - a(n)$, and $|S_1(2)| = 0$. We prove the following.

Lemma 3.2.12. *Let $\omega = \omega(n) : \mathbb{N} \to \mathbb{R}$ be such that $\omega(n) \to \infty$ as $n \to \infty$ and $n^{1/3}p \geq \omega$. Suppose that $a(n) : \mathbb{N} \to \mathbb{R}$ is such that $a(n) = \Omega(np^{1/2})$ but $a(n) \leq \frac{n}{2} - \frac{np}{\omega^{1/3}}$. For any*

$\varepsilon > 0$ *a.a.s. the following holds: for any partition* (P_0, P_1) *of* V_n *with* $|P_0| = n/2 + a(n)$*, all but fewer than* $np/\omega^{1/2}$ *vertices* $v \in V_n$ *we have for* $i \in \{0, 1\}$,

$$|d_{P_i}(v) - |P_i|p| \leq \varepsilon a(n)p.$$

Proof. Let $S \subset V_n$ be such that $|S| = np/\sqrt{\omega}$ and let (P_0, P_1) be a partition of V_n as specified above. We apply the Chernoff bound to deduce the concentration of the number of edges within S. Namely the quantity $e(S)$:

$$\mathbb{P}\left(\left|e(S) - \binom{|S|}{2}p\right| \geq \frac{p|S|^2}{\sqrt{\omega}}\right) \leq \exp\left(-\Omega\left(\frac{p|S|^2}{\omega}\right)\right).$$

As $p \geq \omega n^{-1/3}$, we deduce that $p|S|^2/\omega = \Omega(n\omega)$. Hence with probability at least $1-2^{-\Omega(\omega \cdot n))}$, we have that $e(S)$ is suitably concentrated about its expected value. Thereby, on this event, the average degree on $G(n, p)[S]$ is

$$\frac{2e(S)}{|S|} \leq p|S|(1 + o(1)).$$

So on the above event and for n sufficiently large, there exists a subset $S' \subset S$ with $|S'| = |S|/2$ such that every vertex $v \in S'$ has $d_S(v) \leq 3|S|p$. We first consider P_0. If $|d_{P_0}(v) - |P_0|p| > \varepsilon a(n)p$, for any $v \in S$, then of course this is also the case for any $v \in S'$. Since $|P_0| \gg |S|$ and, moreover, $a(n) = \Omega(np^{1/2}) \gg |S|$, it follows that for any $v \in S'$

$$|d_{P_0 \setminus S}(v) - |P_0 \setminus S|p| > \varepsilon a(n)p - 3|S|p > \frac{\varepsilon}{2}a(n)p,$$

provided that n is sufficiently large. Observe that the family $\{d_{P_0 \setminus S}(v)\}_{v \in S'}$ consists of i.i.d. random variables which are binomially distributed with parameters $|P_0 \setminus S|, p$. The Chernoff bound implies that

$$\mathbb{P}\left[\text{for all } v \in S' \ |d_{P_0 \setminus S}(v) - |P_0 \setminus S|p| > \frac{\varepsilon}{2}a(n)p\right] = \exp\left(-\Omega\left(\frac{|S'|a(n)^2p^2}{np}\right)\right).$$

But since $|S'| = |S|/2$ and $|S| = np/\sqrt{\omega}$ we have

$$\frac{|S'|a(n)^2p^2}{np} = \Omega\left(\frac{a^2(n)p^2}{\sqrt{\omega}}\right).$$

Since $a(n) = \Omega(np^{1/2})$, the numerator on the right hand side is

$$a(n)^2p^2 = \Omega(n^2p^3).$$

Hence as $p \geq \omega n^{-1/3}\omega$ we have that $a(n)^2p^2/\sqrt{\omega} \geq n\omega^{5/2}$, thus this probability is $o(e^{-n})$. If instead we consider $d_{P_1}(v)$, the calculations remain almost identical; though we remark that as $|P_0| \leq n - np/\omega^{1/3}$ we have that $|P_1| \geq np/\omega^{1/3} \gg |S|$. Thus the calculation of the appropriate probability follows with:

$$\mathbb{P}\left[\left|d_{P_1\setminus S}(v) - |P_1 \setminus S|p\right| > \frac{\varepsilon}{2}a(n)p, \text{ for all } v \in S'\right] = \\ \exp\left(-\Omega\left(\frac{\omega|S'|a(n)^2p^2}{np}\right)\right) = \exp\left(-\Omega(n\omega^{7/2})\right).$$

As there are $2^{O(n)}$ choices for S and the partition (P_0, P_1), the union bound implies the lemma. $\square$

We now apply Lemma 3.2.12 to deduce the evolution of the behaviour of all but at most $np/\omega^{1/2}$ vertices from round 1 into round 2. By our previous arguments, and potentially permuting the rows of Q and relabelling the vertices, we have $|S_1(0)| = n/2 + a(n)$ and $|S_1(1)| = n/2 - a(n)$, with $|S_1(2)| = o(1)$. We define Q' to be the 3×2 matrix, which is formed by removing column 2 from Q. We again distinguish cases on Q' depending on its row sums.

In light of Lemma 3.2.12, let X be the set of vertices which satisfy the concentration of degrees inside $S_1(0)$ and $S_1(1)$, i.e for an $a(n)$ satisfying the hypothesis of Lemma 3.2.12, X consists of vertices satisfying $|d_{P_i}(v) - |P_i|p| \leq \varepsilon a(n)p$ for $i \in \{0, 1\}$. By Lemma 3.2.12 we have that $|X| \geq n - np/\omega^{1/2}$ a.a.s. We claim that for vertices in X their evolution is unanimous and entirely determined by the entries in Q'.

Lemma 3.2.13. *Suppose that after one round we have that* $|S_1(0)| = n/2 + a(n)$ *and* $|S_1(1)| = n/2 - a(n)$. *Let* $\Sigma R'_\ell = q_{\ell,0} + q_{\ell,1}$ *denote the sum of row* ℓ *in the matrix* Q'. *Then for any vertex* $v \in X$ *we have the following for* $i, i^* \in \{0, 1, 2\}$,

$$\operatorname{sgn}\left(T(v;i) - T(v;i^*)\right) = \operatorname{sgn}\left(\Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*}\beta_{\mathbf{c}^T\cdot\mathbf{C}}\right).$$

Proof. We consider the decomposition of the payoff difference $T_1(v;i) - T_1(v;i^*)$. We use the notation $x = a \pm b$ to denote that $a - b \le x \le a + b$, for $b \ge 0$. As $v \in X$, by Lemma 3.2.12 for any $\varepsilon > 0$ we have that for $k \in \{0, 1\}$ that $n_1(v;k) = n_k p \pm \varepsilon a(n)p$. Hence we have that $n_1(v;0) = p(n/2 + a(n)) \pm \varepsilon a(n)p$, and $n_1(v;1) = p(n/2 - a(n)) \pm \varepsilon a(n)p$. Furthermore, we recall by equation (3.25) that

$$\frac{n}{2} + n\beta_{\mathbf{c}^T\cdot\mathbf{C}} - \frac{\delta}{3}np^{1/2} \le X_0 \le \frac{n}{2} + n\beta_{\mathbf{c}^T\cdot\mathbf{C}} + \frac{\delta}{3}np^{1/2}. \tag{3.40}$$

Hence as $X_0 = n_0$ we may utilise the above notation to deduce that,

$$a(n) = n\beta_{\mathbf{c}^T\cdot\mathbf{C}} \pm \frac{\delta}{3}np^{1/2}.$$

In the following calculations, we may take ε and δ to be chosen sufficiently small for their associated terms to have negligible effect when n is large. We give an explicit bound further within the proof. By expanding the payoff we can choose some K_1, K_2 large enough such that

$$\begin{aligned} T(v;i) - T(v;i^*) &= (q_{i,0} - q_{i^*,0})\left(\frac{np}{2} + a(n)p \pm \varepsilon a(n)p\right) + (q_{i,1} - q_{i^*,1})\left(\frac{np}{2} - a(n)p \pm \varepsilon a(n)p\right) \\ &= \frac{np}{2}\left(\Sigma R'_i - \Sigma R'_{i^*}\right) + a(n)pC^{0,1}_{i,i^*} \pm K_1\varepsilon a(n)p \\ &= \frac{np}{2}\left(\Sigma R'_i - \Sigma R'_{i^*} + C^{0,1}_{i,i^*}\beta_{\mathbf{c}\cdot\mathbf{C}}\right) \pm \left(K_1\varepsilon a(n)p + \frac{K_2\delta}{3}np^{3/2}\right) \end{aligned} \tag{3.41}$$

We observe we may choose δ and ε such that,

$$\delta, \varepsilon < \frac{1}{100}(\Sigma R'_i - \Sigma R'_{i^*} + C^{0,1}_{i,i^*}\beta_{\mathbf{c}\cdot\mathbf{C}}) \cdot \min\left\{\frac{1}{K_1}, \frac{1}{K_2}, \frac{1}{C^{0,1}_{i,i^*}}\right\}.$$

Consequently, this implies that the sign of equation (3.41) is decided by the leading term. Thus for n sufficiently large, we have,

$$\operatorname{sgn}\left(T(v;i) - T(v;i^*)\right) = \operatorname{sgn}\left(\Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*}\beta_{\mathbf{c}^T\cdot\mathbf{C}}\right). \qquad \square$$

The above deciding term is additive in the sense of Lemma 3.1.1, thus given that each term is non-zero, a preferred strategy is decided. The following three results justify that these terms are non-zero with high probability. The first lemma and corollary imply that rows with equal row sums in Q', either have a non-zero cost coefficient or they are dominated column-wise by the remaining row. This causes all vertices to prefer the dominating strategy deterministically. While our final lemma implies that the probability $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$ takes a value near to

$$\beta^{(i,i^*)} := -\frac{\Sigma R'_i - \Sigma R'_{i^*}}{2C^{0,1}_{i,i^*}},$$

for any $i, i^* \in \{0,1,2\}$, is at most ε for any $\varepsilon > 0$. Together these results imply that,

$$\Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*}\beta_{\mathbf{c}^T\cdot\mathbf{C}} \neq 0$$

with probability at least $1 - \varepsilon$.

Lemma 3.2.14. *Suppose Q' is the 3×2 matrix formed by removing column 0 from Q. If $\Sigma R'_i = \Sigma R'_{i^*}$ and they are maximal across the row sums in Q', then we have that $C^{0,1}_{i,i^*} \neq 0$. Furthermore, if $\Sigma R'_i = \Sigma R'_{i^*} < \Sigma R'_{i^{**}}$ and $C^{0,1}_{i,i^*} = 0$, then both column maxima occur in row i^{**}.*

Proof. Suppose that $\Sigma R'_i = \Sigma R'_{i^*}$. If $C^{0,1}_{i,i^*} = 0$, then it follows that $q_{i,0} - q_{i,1} - q_{i^*,0} + q_{i^*,1} = 0$. However as rows i and i^* have the same row sum, it then follows that $q_{i,0} + q_{i,1} = q_{i^*,0} + q_{i^*,1}$.

Thus by solving the resultant system, we conclude that $q_{i,0} = q_{i^*,0}$ and $q_{i,1} = q_{i^*,1}$.

But we recall that Q, and hence Q', have unique column maxima. From the above, we observe that row i and i^* are identical and $\Sigma R'_i = \Sigma R'_{i^*}$ is strictly greater than the sum of the elements in the remaining row, say i^{**}. Hence it follows that both column maxima in Q' can not occur row i^{**}. Hence there exists a column in Q' with non-unique column maxima, this is a contradiction and therefore $C^{0,1}_{i,i^*} \neq 0$. For the second part, we observe that if $C^{0,1}_{i,i^*} = 0$, then as row i and i^* are identical, both column maxima must occur in row i^{**}. □

Our final lemma is concerned with the setting where $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$ takes a specific value which causes the leading term in equation (3.41) to equal zero for some $i, i^* \in \{0,1,2\}$. For each pair i, i^* the equation,

$$\Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*}\beta_{\mathbf{c}^T\cdot\mathbf{C}} = 0$$

is linear and has a solution given by,

$$\beta^{(i,i^*)} := -\frac{\Sigma R'_i - \Sigma R'_{i^*}}{2C^{0,1}_{i,i^*}},$$

In the following lemma, we show that the probability $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$ takes a value sufficiently close to any fixed value is small, hence the probability that $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$ takes a value close to $\beta^{(i,i^*)}$ for any $i, i^* \in \{0,1,2\}$ is at most ε for any $\varepsilon > 0$.

Lemma 3.2.15. *Let $\beta_{\mathbf{c}^T\cdot\mathbf{C}}$ be defined as above, i.e.*

$$\beta_{\mathbf{c}^T\cdot\mathbf{C}} = \int\limits_{\substack{0>\mathbf{x}^T\cdot\mathbf{C}>-p^{1/2}\mathbf{c}^T\cdot\mathbf{C}+o(p^{1/2}) \\ \|\mathbf{x}\|_\infty<\log n}} \phi(\mathbf{x})\mathrm{d}\mathbf{x}.$$

Suppose $\beta^* \in (0,1/2]$. Then we have that for every $\varepsilon > 0$ there exists a $\xi > 0$ such that

$$\mathbb{P}[|\beta_{\mathbf{c}^T\cdot C} - \beta^*| < \xi] < \varepsilon.$$

Proof. Fix $\beta^* \in (0, 1/2]$ and we set I to be the region integrated over above,

$$I = \{\mathbf{x} = (x_0, x_1) : \|\mathbf{x}\|_\infty \leq \log n,\ 0 > \mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\mathbf{c}^T \cdot \mathbf{C} + o(p^{1/2})\},$$

We view the value of $\beta_{\mathbf{c}^T \cdot \mathbf{C}}$ to be a function of I. In this setting we have conditioned on the event that $\{\mathbf{c}^T \cdot \mathbf{C} > 0\}$. Thus we parameterise based on the value of $\mathbf{c}^T \cdot \mathbf{C}$. For $\lambda > 0$ we define,

$$I_\lambda = \{(x_0, x_1) : \|\mathbf{x}\|_\infty \leq \log n,\ 0 > \mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\lambda + o(p^{1/2})\},$$

where the $o(p^{1/2})$ is identical to that in the definition of I. Hence we have that $I = I_{\mathbf{c}^T \cdot \mathbf{C}}$. We remark that the $o(p^{1/2})$ term contains an implicit dependence on λ, thus we uniformly bound I by above and below. To this end, we define

$$I^*_\lambda = \{(x_0, x_1) : \|\mathbf{x}\|_\infty \leq \log n,\ 0 > \mathbf{x}^T \cdot \mathbf{C} > -p^{1/2}\lambda\}$$

We remark that by equation (3.15) the above $o(p^{1/2})$ term is contributed by a measure of lattice points. Thus it is non-negative and therefore $I_\lambda \subseteq I^*_\lambda$. Similarly by Equation (3.15) we note that this error term is of order at most,

$$O\left(\frac{n(v)^{1/2}}{n} + \frac{1}{n(v)^{1/2}}\right).$$

However we observe for n sufficiently large that,

$$\frac{n(v)^{1/2}}{n} + \frac{1}{n(v)^{1/2}} \leq \frac{2}{\sqrt{np}} \leq \frac{p^{1/2}}{n^{1/6}}.$$

Hence for any such n, $I^*_{\lambda - \frac{2}{n^{1/6}}} \subseteq I_\lambda$. We now define the following integrals over these regions as a function of λ,

$$g(\lambda) = \int_{I_\lambda} \phi(\mathbf{x}) d\mathbf{x},$$

$$g^*(\lambda) = \int_{I^*_\lambda} \phi(\mathbf{x})d\mathbf{x}.$$

By combining the above observations, we note that for any $\lambda > 0$ and n sufficiently large, we have that

$$g^*(\lambda - 2n^{-1/6}) \leq g(\lambda) \leq g^*(\lambda).$$

Furthermore we note that,

$$\begin{aligned} g^*(\lambda) - g^*(\lambda - 2n^{-1/6}) &= \int_{I^*_\lambda} \phi(\mathbf{x})d\mathbf{x} - \int_{I^*_{\lambda - 2n^{-1/6}}} \phi(\mathbf{x})d\mathbf{x} \\ &\leq \left| I^*_\lambda \setminus I^*_{\lambda - 2n^{-1/6}} \right| \\ &= O\left(\frac{\log n}{n^{1/6}}\right). \end{aligned}$$

Now suppose for a given $\beta^* \in (0, 1/2]$ and some $\xi > 0$, the parameter λ is such that,

$$\left| \int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} - \beta^* \right| < \xi.$$

We relate the above as a sub-event about $g^*(\lambda)$. Firstly, by considering the upper tail we have,

$$\int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} < \xi + \beta^*,$$

and therefore, $g^*(\lambda - 2n^{-1/6}) < \xi + \beta^*$. But this is equivalent to

$$g^*(\lambda - 2n^{-1/6}) + g^*(\lambda) - g^*(\lambda) < \beta^* + \xi.$$

Therefore we have,

$$g^*(\lambda) < \beta^* + \xi + \frac{\log n}{n^{1/6}},$$

and we conclude that for n sufficiently large,

$$\mathbb{P}\left[\int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} < \xi + \beta^*\right] \leq \mathbb{P}\left[g^*(\lambda) < 2\xi + \beta^*\right].$$

Similarly, the lower tail is given,

$$\int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} > \beta^* - \xi$$

Then as $g^*(\lambda) > g(\lambda)$, for n large, we have that,

$$\mathbb{P}\left[\int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} > \beta^* - \xi\right] \leq \mathbb{P}\left[g^*(\lambda) > \beta^* - 2\xi\right]$$

Thus it suffices for us to show that for any $\varepsilon > 0$ there exists a $\xi' > 0$ such that,

$$\mathbb{P}\left[|g^*(\lambda) - \beta^*| < \xi'\right] < \varepsilon.$$

We observe that $g^*(\lambda)$ is continuous and monotonically increasing in λ. Furthermore, we have that $g^*(0) = 0$, and $\lim_{\lambda\to\infty} g^*(\lambda) = 1/2$. Hence by the Intermediate Value Theorem, there exists a positive λ^* such that $g^*(\lambda^*) = \beta^*$. Thus by monotonicity, we have that for every $\xi' > 0$ if $|g^*(\lambda) - \beta^*| < \xi'$ then there exists a maximal $\delta_{\xi'}$, such that $|\lambda - \lambda^*| < \delta'_\xi$. Hence we have that,

$$\mathbb{P}\left[\left|\int_{I_\lambda} \phi(\mathbf{x})d\mathbf{x} - \beta^*\right| < \xi\right] \leq \mathbb{P}[|\lambda - \lambda^*| < \delta_{\xi'}].$$

We recall that the above is uniform over for any choice of λ, thereby we set $\lambda = \mathbf{c}^T \cdot \mathbf{C}$. We also observe that the vector c comprises of the scaled deviations of $|S_0(0)|$ and $|S_0(1)|$. By the Central Limit Theorem, it follows that $\mathbf{c}$ converges in distribution to a standard bivariate normal. Equivalently, there exists constants μ and σ such that as $n \to \infty$,

$$\mathbf{c}^T \cdot \mathbf{C} \xrightarrow{d} N \sim \mathrm{N}(\mu, \sigma^2)$$

Hence for n large enough we have that,

$$\mathbb{P}[|\lambda - \lambda^*| < \delta'_\xi] < \mathbb{P}[|N - \lambda^*| < \delta_{\xi'}] + \frac{\varepsilon}{2}.$$

Thus we have that,

$$\mathbb{P}[|N - \lambda^*| < \delta_{\xi'}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{\lambda^* - \delta'_\xi}^{\lambda^* + \delta_{\xi'}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \mathrm{d}x \leq \frac{2\delta_{\xi'}}{\sigma\sqrt{2\pi}}.$$

By Lemma 3.4.3, we have that uniformly over any choice of $\delta_{\xi'}$, it follows that $\delta_{\xi'} \to 0$ as $\xi' \to 0$. Therefore, by taking ξ sufficiently close to zero, we have that,

$$\mathbb{P}[|N - \lambda^*| < \delta_{\xi'}] < \frac{\varepsilon}{2},$$

and thus it follows that,

$$\mathbb{P}|\beta_{\mathbf{c}^T \cdot \mathbf{C}} - \beta^*| < \xi'] < \varepsilon.$$

□

Hence by the union bound, the probability $\beta_{\mathbf{c}^T \cdot \mathbf{C}}$ is within a small interval containing one of the following: $\beta^{(0,1)}, \beta^{(1,2)}$ or $\beta^{(0,2)}$ is at most ε. Thus the following lemma determines the evolution of vertices in X.

Lemma 3.2.16. *Suppose that $n_0 = n/2 + a(n)$ and $n_1 = n/2 - a(n)$, with $a(n) > \frac{\delta}{2} np^{1/2}$. Let X be the set of vertices which satisfy the conclusion of Lemma 3.2.12. Then there exists a unique strategy i^*, such that for all $v \in X$, we have that v switches to strategy i^* in the next round.*

Proof. We split this proof into three cases, depending on how many rows share the same maximum partial row sum. We recall by Lemma 3.2.13 that,

$$\operatorname{sgn}\left(T(v; i) - T(v; i^*)\right) = \operatorname{sgn}\left(\Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*} \beta_{\mathbf{c}^T \cdot \mathbf{C}}\right).$$

For $i, i^* \in \{0, 1, 2\}$ we set

$$T_{i,i^*} = \Sigma R'_i - \Sigma R'_{i^*} + 2C^{0,1}_{i,i^*} \beta_{\mathbf{c}^T \cdot \mathbf{C}}.$$

Immediately we observe a similar identity to that of Lemma 3.1.1 namely,

$$T_{0,1} + T_{1,2} = T_{0,2}. \tag{3.42}$$

Furthermore recall if $T_{i,i^*} > 0$ then $T_1(v; i) > T_1(v; i^*)$, if this event occurs then we say that $i \succ i^*$. Similarly if $T_{i,i^*} < 0$ then $i^* \succ i$. We now split into a number of cases, depending on the values of $\Sigma R'_i$ for $i \in \{0, 1, 2\}$.

Firstly, if all three partial row sums in Q' are distinct, then by Lemma 3.2.15 all T_{i,i^*} terms are non-zero with high probability. Thus by the above discussion, the T_{i,i^*} terms determine a winner. A winner is found, as the additive property in (3.42) prevents the existence of the two possible *Condorcet cycles*, (i.e $0 \succ 1 \succ 2 \succ 0$ and $0 \succ 2 \succ 1 \succ 0$).

Suppose now that there are exactly two partial row sums that are the same, say in row i and i^*, while i^{**} denotes the distinct partial row sum. Then Lemma 3.2.15 implies that $T_{i,i^{**}}$ and $T_{i^*,i^{**}}$ are non-zero with high probability. We split into two cases. If $\Sigma R'_i = \Sigma R'_{i^*} > \Sigma R'_{i^{**}}$ then Lemma 3.2.14 implies that T_{i,i^*} is also non-zero, and we are done. If we have that $\Sigma R'_{i^{**}} > \Sigma R'_i = \Sigma R'_{i^*}$, and $T_{i,i^*} \neq 0$ then we are done. Otherwise if $T_{i,i^*} = 0$, then this implies that $C^{0,1}_{i,i^*} = 0$, and by Lemma 3.2.14 we have that both column maxima occur in row i^{**}. Thus this row dominates column-wise and all vertices in V_n will switch to i^{**}.

Finally suppose we have that $\Sigma R'_i = \Sigma R'_{i^*} = \Sigma R'_{i^{**}}$. By Lemma 3.2.14, all associated cost coefficients are non-zero. This implies that all $T_{i,j}$ terms are non-zero, and thus a winner is decided as above. □

3.2.5 The Final Round

Following the previous section, we detail the final argument to declare unanimity. We recall that the vertices are partitioned into sets $\{S_2(0), S_2(1), S_2(2)\}$. Furthermore, we are in the setting where for some $i^* \in \{0, 1, 2\}$ there exists a subset $P_{i^*} \subset S_2(i^*)$, of size at least $n - np/\omega^{1/2}$. The remaining vertices in $V_n \setminus P_{i^*}$ may be playing any of the three strategies. By relabelling vertices and permuting the corresponding rows of Q we may assume without loss of generality that $i^* = 0$.

We claim that the evolution of the system is now decided by the zero column. This follows from the fact that vertex degrees inside P_0 are asymptotically larger than in $V_n \backslash P_0$. Consequently we have,

$$n_2(v; 1), n_2(v; 2) = o(n_2(v; 0)).$$

As a result, we have,

$$\text{sgn}(T_2(v; i) - T_2(v; j)) = \text{sgn}(q_{i,0} - q_{j,0}).$$

Consequently, as all columns in Q have a unique column maximum, this implies that all vertices switch to the row containing the column maximum of column 0. We recall the event $\mathcal{N}_v$ which holds when $|n(v) - np| < \log n(np)^{1/2}$. As previously remarked this event holds with probability $1 - o(1/n)$. By the union bound, the event $\mathcal{N}' = \cap_{v \in V_n} \mathcal{N}_v$, holds with probability $1 - o(1)$.

Lemma 3.2.17. *Let $P_0 \subset V_n$ be a set of vertices as above, $\omega(n) \to \infty$ and $n \to \infty$. Suppose Q is a 3×3 payoff matrix with unique column maxima and suppose that $|P_0| \geq n - np/\omega^{1/2}$. Then conditional on $\mathcal{N}'$ we have for all $v \in V_n$,*

$$S_3(v) = \underset{0 \leq i \leq 2}{\text{argmax}} \{q_{i,0}\}$$

Proof. Suppose $v \in V_n$ and the event $\mathcal{N}'$, and hence $\mathcal{N}_v$, has occurred. We lead by considering bounds on the number of vertices playing each strategy in the neighbourhood of v. As stated above, all vertices in P_0 are in the 0 state. Therefore we have that, $n_2(v; 1), n_2(v; 2) \leq np/\omega^{1/2}$. Now we observe that as we are conditional on $\mathcal{N}_v$ it follows that, $d_{V_n}(v) > np - \log n(np)^{1/2}$. However we have that $d_{V_n}(v) = d_{P_0}(v) + d_{P_0^c}(v)$ hence we also have that,

$$n_2(v;0) > d_{P_0}(v) > np - \log n(np)^{1/2} - \frac{np}{\omega^{1/2}} = (1-o(1))np.$$

Therefore this implies that $n_2(v;1)$ and $n_2(v;2) = o(n_2(v;0))$. We now observe the payoff difference between two strategies $i, j \in \{0,1,2\}$,

$$T_2(v;i) - T_2(v;j) = \sum_{k=0}^{2} n_2(v;k)(q_{i,k} - q_{j,k}) = n_2(v;0)(q_{i,0} - q_{j,0}) + o(np).$$

Hence, for n sufficiently large

$$\operatorname{sgn}(T_2(v;i) - T_2(v;j)) = \operatorname{sgn}(q_{i,0} - q_{j,0}).$$

Thus it follows that v will switch to the strategy given by the $\operatorname{argmax}_{0\le i\le 2}\{q_{i,0}\}$ in the next round. □

Therefore, all vertices in V_n switch to the $\operatorname{argmax}_{0\le i\le 2}\{q_{i,0}\}$ in round 3, and hence the system enters unanimity, with behaviour dictated by Lemma 1.2.3.

3.3 Discussion

In this chapter, we have considered the evolution of best response dynamics for systems with more than two strategies. We have deduced a classification for payoff matrices Q which differentiate how the system evolves. We observe that our $M(Q)$ quantity mirrors the behaviour of λ in the two strategy case, i.e., $M(Q) = 1$ reflects the case where $\lambda \neq 1$ within the two-strategy system. In the case $M(Q) = 1$ a direct calculation, based on the concentration of the degrees, deduces unanimity in the first round, we observe parallels with Theorem 2.2.10.

The case for $M(Q) = 2$ is significantly more involved and requires the derivation of a new local limit theorem, alongside careful approximations to the normal distribution.

In comparison with the above, we consider $p \gg n^{-1/3}$. Again we show rapid unanimity within at most three rounds, with high probability.

A natural extension is to classify the system for all payoff matrices for values of $2 < M(Q) \leq \ell$. One notable case is for $\ell = 3$ and $Q = I_3$, the 3×3 identify matrix. Deriving the evolution rules in this case leads to a system evolving with the rules given by majority dynamics; analogously $Q = -I_3$ gives rise to minority dynamics. Clearly, $M(I) = 3$, thus the case $M(Q) = \ell$ can give rise to generalisations of majority and minority dynamics.

The main challenge is to determine the bias gained by each strategy from the matrix. For $M(Q) = 2$ the bias gained is given by a function of $\mathbf{c}^T \cdot \mathbf{C}$, a term containing both the randomness of the initial distribution, and values from the matrix. Geometrically this describes a gain in the area within $\mathcal{L}'$ for one strategy. We observe that majority and minority dynamics operate without the presence of a payoff matrix, hence the initial bias is entirely decided by the random initial distribution. A more careful case distinction may be needed here to deal with both the presence and absence of a bias given by the matrix.

3.4 Auxiliary Results for $M(Q) = 2$

In this short, appendix like, section we provide a selection of three fairly straightforward lemmas, which support our analysis for the case $M(Q) = 2$.

3.4.1 Proof That $\mathbf{c}^T \cdot \mathbf{C}$ is Non-Zero

We recall the vectors $\mathbf{c} = (c_0, c_1)^T$ and $\mathbf{C} = \left(C_{0,1}^{0,2}, C_{0,1}^{1,2}\right)^T$, we justify that $\mathbf{c}^T \cdot \mathbf{C}$ is non-zero with high probability.

Lemma 3.4.1. *Suppose Q is a payoff matrix with $M(Q) = 2$, and the maximum row sums occur in rows 0 and 1. For every $\varepsilon > 0$ there exists a $\delta > 0$ such that for all n*

sufficiently large we have,

$$\mathbb{P}\left[|\mathbf{c}^T \cdot \mathbf{C}| < \delta\right] < \varepsilon.$$

Proof. We observe that in the case of $\mathbf{C} \neq 0$, the above holds as a direct consequence of the Central Limit Theorem, using a similar argument as in Lemma 3.2.15. Suppose for a contradiction that $\mathbf{C} = \mathbf{0}$, then if follows that $C_{0,1}^{0,2} = C_{0,1}^{1,2} = 0$. Furthermore, we also have that $\Sigma R_0 = \Sigma R_1$. Hence these equations form the following system,

$$q_{0,0} - q_{1,0} = q_{0,2} - q_{1,2}$$
$$q_{1,1} - q_{0,1} = q_{1,2} - q_{0,2}$$
$$q_{0,0} + q_{0,1} + q_{0,2} = q_{1,0} + q_{1,1} + q_{1,2}$$

Combining the first two lines and re-arranging the bottom we observe,

$$q_{0,0} - q_{1,0} = q_{0,2} - q_{1,2} = -(q_{1,1} - q_{0,1})$$
$$(q_{0,0} - q_{1,0}) + (q_{0,2} - q_{1,2}) = q_{1,1} - q_{0,1}$$

As a result, we observe that,

$$q_{0,0} = q_{1,0},\ q_{0,2} = q_{1,2},\ q_{1,1} = q_{0,1}$$

Hence it follows that rows 0 and 1 are identical. Thus in order to preserve unique column maxima, all column maxima are in row 2, but this contradicts the fact that rows 0 and 1 have strictly maximal row sums. □

3.4.2 The Integral of $\phi(\mathbf{x})$ Outside of $\mathcal{B}^{\infty}_{\log n}$

We briefly compute the integral of $\phi(\mathbf{x})$ over the region $\|\mathbf{x}\|_\infty > \log n$.

Lemma 3.4.2. *For $\phi(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{R}$ defined as $\phi(\mathbf{x}) = \frac{1}{2\pi}|\Sigma|^{1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma\mathbf{x}\right)$ we have,*

$$\int\limits_{\|\mathbf{x}\|_\infty > \log n} \phi(\mathbf{x})\mathrm{d}\mathbf{x} = O\left(e^{-\log^2 n}\right).$$

Proof. We recall that $\phi(\mathbf{x}) = \frac{1}{2\pi}|\Sigma|^{1/2} \exp\left(-\frac{1}{2}\mathbf{x}^T\Sigma\mathbf{x}\right)$, where Σ is positive-definite and symmetric. Hence there exists square matrices P and Λ such that,

$$\Sigma = P\Lambda P^{-1}$$

where P is such that $P^T = P^{-1}$, and Λ is a diagonal matrix containing the eigenvalues of Σ. Thus we may make a change of coordinates by $\mathbf{z} = (z_0, z_1) = P^T\mathbf{x}$. We first consider an upper bound on the transformed region of integration. Namely for any $x \in \mathbb{R}^2$ with $\|\mathbf{x}\|_\infty \geq \log n$ has $\|P^T\mathbf{x}\|_\infty \geq \frac{1}{\sqrt{2}}\log n$. Indeed, if $\|\mathbf{x}\|_\infty \geq \log n$, then $\|\mathbf{x}\|_2 \geq \log n$ as well. But

$$\|\mathbf{z}\|_2 = \mathbf{z}^T\mathbf{z} = \mathbf{x}PP^T\mathbf{x} = \mathbf{x}\mathbf{x} = \|\mathbf{x}\|_2.$$

So $\|\mathbf{z}\|_2 \geq \log n$ as well. But then $\|\mathbf{z}\|_2^2 \geq \log^2 n$ and, therefore, $\max\{z_0^2, z_1^2\} \geq \frac{1}{2}\log^2 n$. Thereby, $\|\mathbf{z}\|_\infty \geq \frac{1}{\sqrt{2}}\log n$. Hence by applying the above and observing that $\det(P) = 1$, we have:

$$\int\limits_{\|\mathbf{x}\|_\infty > \log n} \phi(\mathbf{x})\mathrm{d}\mathbf{x} \leq \frac{1}{2\pi}|\Sigma|^{1/2} \int\limits_{\|\mathbf{z}\|_\infty > \frac{1}{\sqrt{2}}\log n} e^{-\frac{1}{2}\mathbf{z}^T\Lambda\mathbf{z}}\mathrm{d}\mathbf{z}.$$

We note that as Λ is diagonal, and contain positive entries λ_0 and λ_1. We have that the above integral is separable into z_0 and z_1.

$$\int\limits_{\|\mathbf{z}\|_\infty > \frac{1}{\sqrt{2}}\log n} e^{-\frac{1}{2}\mathbf{z}^T\Lambda\mathbf{z}} \leq 2\int_{-\infty}^{\infty} e^{-\frac{1}{2}\lambda_0 z_0^2}\mathrm{d}z_0 \int_{\frac{1}{\sqrt{2}}\log n}^{\infty} e^{-\frac{1}{2}\lambda_1 z_1^2}\mathrm{d}z_1 +$$
$$2\int_{-\infty}^{\infty} e^{-\frac{1}{2}\lambda_1 z_1^2}\mathrm{d}z_1 \int_{\frac{1}{\sqrt{2}}\log n}^{\infty} e^{-\frac{1}{2}\lambda_0 z_0^2}\mathrm{d}z_0.$$

As Σ is positive-definite we have that $\lambda_0, \lambda_1 > 0$, furthermore we observe by standard computation that for all $a > 0$,

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}at^2} \mathrm{d}t = O(1).$$

Finally we observe for $a > 0$,

$$\int_{\log n}^{\infty} e^{-\frac{1}{2}at^2} \mathrm{d}t = \frac{1}{\sqrt{a}} \cdot \mathrm{erfc}(\log n) = O\left(e^{-\log^2 n}\right).$$

where erfc is the complementary error function [87]. □

3.4.3 A Lemma Concerning Continuous Monotone Functions

We briefly state a lemma concerning the property of a function from $\mathbb{R} \to \mathbb{R}$ which is continuous and strictly monotone. We consider the standard $\varepsilon - \delta$ definition of continuity and further show that under the assumption of monotonicity, we have that $\delta \to 0$ as $\varepsilon \to 0$.

Lemma 3.4.3. *Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and strictly monotone (increasing) on $\mathbb{R}$. Let $x_0 \in \mathbb{R}$, then for every $\varepsilon > 0$ there exists $\delta > 0$ such that,*

$$|f(x) - f(x_0)| < \varepsilon \text{ whenever } |x - x_0| < \delta,$$

Moreover, for each $\varepsilon > 0$, let δ_ε be any value of δ which satisfies the definition above. Then for any choice of suitable values for the function δ_ε, we have that $\delta_\varepsilon \to 0$ as $\varepsilon \to 0$.

Proof. We note that for any $\varepsilon > 0$ we have that at least one choice of δ_ε trivially exists by the standard definition of continuity. Thus we show that for any $\hat{\delta} > 0$ we have that there exists an $\hat{\varepsilon}$ such that $\delta_\varepsilon < \hat{\delta}$ whenever $\varepsilon < \hat{\varepsilon}$. We set,

$$\hat{\varepsilon} = \min\left\{\frac{f(x_0 + \hat{\delta}) - f(x_0)}{2}, \frac{f(x_0) - f(x_0 - \hat{\delta})}{2}\right\},$$

by monotonicity we have that $\hat{\varepsilon} > 0$. Similarly by monotonicity we also have that

$$f(x_0) < f(x_0) + \hat{\varepsilon} \leq \frac{f(x_0 + \hat{\delta}) + f(x_0)}{2} < f(x_0 + \hat{\delta}).$$

Thus by the Intermediate Value Theorem there exists some $\delta^+ > 0$ such that, $f(x_0 + \delta^+) = f(x_0) + \hat{\varepsilon}$, and hence by monotonicity we have that $\delta^+ < \hat{\delta}$. Similarly we also have that,

$$f(x_0) > f(x_0) - \hat{\varepsilon} > f(x_0 - \hat{\delta}).$$

Following from this, we can analogously define δ^- and again conclude that $\delta^- < \hat{\delta}$. Now by the definition of continuity at x_0 we have for any $\varepsilon < \hat{\varepsilon}$ that there exists δ_ε such that,

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$$

whenever $|x - x_0| < \delta_\varepsilon$. But $f(x_0) + \varepsilon < f(x_0 + \delta^+)$ and $f(x_0) - \varepsilon > f(x_0 - \delta^-)$. Hence by monotonicity for any suitable choice of δ_ε we have that $\delta_\varepsilon \leq \max\{\delta^-, \delta^+\} < \hat{\delta}$, this choice was uniform over all such choice of δ_ε hence the result follows. $\square$

CHAPTER 4

THE MODULARITY OF HYPERBOLIC RANDOM GRAPHS ON THE PLANE

4.1 Introduction

In this chapter, we consider the problem of computing the modularity score of the KPKBV Hyperbolic Random Graph Model. The definition of this model (and its poissonisation) can be found in Section 1.3.3, while the definition of the modularity partition function can be found in Section 1.3.2. We lead this section with a survey of the current literature on the KPKBV model and a general discussion of recent work in graph modularity. We close this section with two preliminary results involving approximating the hyperbolic ball around a point, and a useful mapping between $\mathcal{D}_R$ and $\mathbb{R}^2$. The remainder of this chapter is dedicated to proving Theorem 1.3.1, namely that the hyperbolic random graph possesses optimal modularity with high probability:

Theorem 1.3.1. *For any $\alpha > 1/2$ and $\nu > 0$, we have*

$$\operatorname{mod}(\mathcal{P}(n; \alpha, \nu)) \to 1,$$

as $n \to \infty$, in probability.

4.1.1 Typical Properties of the KPKBV Model

For $\alpha \in (1/2, \infty)$, Krioukov et al. [50] show that the tails of the distribution of the degrees in $\mathcal{G}(n; \alpha, \nu)$ follow a power law with exponent $2\alpha + 1$. This was verified rigorously by Gugelman et al. in [41]. Thus, when $\alpha \in (1/2, 1)$ the exponent is between 2 and 3. There has been experimental evidence that this is indeed the case in a number of networks that emerge in applications (the survey [3] contains a comprehensive list of such examples). Krioukov et al. [50] also observe that the average degree of $\mathcal{G}(n; \alpha, \nu)$ is also tuned by the parameter ν for $\alpha \in (1/2, \infty)$. This was also proved by Gugelman et al. [41]. They showed that the average degree tends to $8\alpha^2\nu/\pi(2\alpha - 1)^2$ in probability. However, when $\alpha \in (0, 1/2]$, the average degree tends to infinity as $n \to \infty$. Thus, in this sense, the regime $\alpha \in (1/2, \infty)$ corresponds to the so-called *thermodynamic regime* in the context of random geometric graphs on the Euclidean plane [76].

Gugelman et al. [41] also showed $\mathcal{G}(n; \alpha, \nu)$ has *clustering coefficient* that is a.a.s. bounded away from 0. More precise results about the scaling of the local clustering coefficient in terms of the degrees of the vertices were obtained by Stegehuis et al. [84]. More recently in [34], convergence in probability of the clustering coefficient to an explicitly determined constant was derived.

When α is small, the density ρ_n induces more points near the origin and one may expect increased graph connectivity there. In [13], Bode et al. proved that $\alpha = 1$ is the critical point for the emergence of a giant component in $\mathcal{G}(n; \alpha, \nu)$. When $\alpha \in (0, 1)$, the fraction of the vertices contained in the largest component is bounded away from 0 a.a.s. [13], whereas if $\alpha \in (1, \infty)$, the largest component is sublinear in n a.a.s. For $\alpha = 1$, the component structure depends on ν. If ν is large enough, then a giant component exists a.a.s., but if ν is small enough, then a.a.s. all components have sublinear size [13].

The above results were strengthened in [33]. In that paper, it was shown that the fraction of vertices which belong to the largest component converges in probability to a certain constant which depends on α and ν. More specifically, when $\alpha = 1$, it turns out that there exists a critical value $\nu_0 \in (0, \infty)$ such that when ν crosses ν_0 a giant component emerges a.a.s. The papers [47] and [48] consider the size of the second largest component. Therein, it is shown that when $\alpha \in (0,1)$ the second largest component has polylogarithmic order a.a.s.

The connectivity of $\mathcal{G}(n;\alpha,\nu)$ was considered by Bode et al. in [14]. There, it is shown that for $\alpha < 1/2$ the random graph $\mathcal{G}(n;\alpha,\nu)$ is a.a.s. connected, it is disconnected for $\alpha > 1/2$. When $\alpha = 1/2$, it turns out that the probability of connectivity converges to a certain constant which is given explicitly in [14].

The a.a.s. disconnectedness of $\mathcal{G}(n;\alpha,\nu)$ for $\alpha > 1/2$ follows easily from the a.a.s. existence of isolated vertices. Recently, asymptotic distributional properties of the number of isolated as well as the extreme points in the poissonisation of $\mathcal{G}(n;\alpha,\nu)$ were derived in [35]. (A point is called extreme, when it is not connected to any other point of a larger radius.) The authors showed that the former satisfies a central limit theorem when $\alpha > 1$, but it does not when $\alpha < 1$. However, the number of extreme points satisfies a central limit theorem for any $\alpha > 1/2$. This is due to the fact that the number of isolated vertices is sensitive to the existence of a few vertices close to the centre of $\mathcal{D}_R$. Those a.a.s. appear when $1/2 < \alpha < 1$. On the other hand, extreme points involve only local dependencies.

Bounds on the diameter of $\mathcal{G}(n;\alpha,\nu)$ were derived in [47] and [36]. Therein, polylogarithmic upper bounds on the diameter are shown. These were improved by Müller and Staps [66] who deduced a logarithmic upper bound on the diameter. Furthermore, in [1] it is shown that for $\alpha \in (1/2,1)$ the largest component has doubly logarithmic typical distances and it forms what is called an *ultra-small world*.

4.1.2 Survey of Results on Graph Modularity

Newman [69] determined the modularity of several examples of complex networks, not only social, finding them ranging between 0.3 and 0.8. Among these examples, higher modularity (> 0.7) was found in the social network of co-authorship among scientists working on condensed matter.

Brandes et al. [16] showed that finding the modularity of a given graph is NP-hard.

Further it was established by Dinh, Li and Thai that it is NP-hard to approximate modularity to within any constant factor [28]. However, community detection in networks has been a central theme in network science. Newman [70] used modularity to design a spectral algorithm for community detection in a given network. A popular algorithm, the *Louvain method*, is an iterative clustering technique using the modularity function to compare candidate partitions [11]. Methods that utilise random walks on the network in order to trace its community structure are the *Infomap algorithm* [82] and the *Walktrap algorithm* [78]. The *Label propagation algorithm* [81] on the other hand uses an approach that is similar to majority dynamics where the states of the nodes are labels that indicate different communities. More recently, Wang et al. [86] came up with another method which uses the notion of *local expansion*. We further note that in [86], the authors give a comparison of these methods on a number of networks. According to Table 3 therein, typically, the modularity of these networks, as it is evaluated by these algorithms, is far from 1 (we see values ranging from 0.39 to 0.94).

For binomial random graphs from the $G(n,p)$ model, there is a transition for the typical behaviour of $\mathrm{mod}(G(n,p))$ that is determined by np. In particular, McDiarmid and Skerman showed in [60], that when $np \leq 1 + o(1)$, then $\mathrm{mod}(G(n,p))$ is concentrated around 1, but when np exceeds and is bounded away from 1, then it scales like $(np)^{-1/2}$. They have also shown [59] that for random d-regular graphs of bounded degree, it is bounded away from 0 and 1 with high probability and scales approximately like $1/\sqrt{d}$ when d is large. Recently, Lichev and Mitsche [56] showed that for $d = 3$ the modularity exceeds $2/3$ (confirming a conjecture of McDiarmid and Skerman) and is below 0.8 with high probability. They further considered the modularity of random graphs having a given degree distribution with a bounded maximum degree.

The relation between the topological properties of a graph and its modularity was explored in [53, 59]. In particular, Lasoń and Sulkowska [53] showed that the class of graphs with an excluded minor and sub-linear maximum degree have modularity that approaches 1 as their number of vertices grows.

In the context of community driven models, Zuev et al. [88] introduced a model (*preferential geometric attachment*) which encapsulates the notion of an underlying hyperbolic space of hierarchies where the angle of each vertex is not selected uniformly at random, but it follows a distribution which is biased towards sectors that are more densely populated. This aims at modelling the notion of *homophily* through a preferential attachment mechanism: a newly arrived node is more likely to be located in a region/community of a larger size. In this model, the notion of *soft communities* is considered in which nodes form a community if they are close to each other in terms of their relative angle. This model rectifies the lack of community structure that is exhibited in the *popularity-similarity-optimisation* model that was introduced by Papadopoulos et al. [75]. Later on Muscoloni and Cannistraci [67] introduced the *non-uniform* popularity-similarity-optimisation model which is also a strengthening of the model of Zuev et al. This model gives the opportunity to fix the number and the size of communities and tune their mixing. Furthermore, the authors use in [4] this model as a benchmark for the algorithms we described above.

On the other hand, the classic *preferential attachment model* of Barabási and Albert [8] has modularity which is bounded away from 1. In particular, Prokhorenkova et al. [80] observed that this is bounded away from 1 with high probability (in the version of the model where each newly arriving vertex is attached to $m \geq 2$ existing vertices - in fact, they show that for $m \geq 3$ the modularity is at most $15/16$ with high probability). This is not the case for the *spatial preferential attachment model* which was introduced by Aiello et al. [2]. This is a geometric version of the preferential attachment model where vertices arrive one at a time at the d-dimensional unit cube and are allocated uniformly therein. They are attached to existing vertices if they are within the *sphere of influence* of the latter. The size of this is proportional to their degree. Ostroumova Prokhorenkova et al. [80] showed that the modularity of the resulting random graph approaches 1 with high probability as the number of vertices tends to infinity.

4.1.3 Preliminaries: Approximating a Ball Around a Point

The main lemma in this section provides a useful (almost) characterisation of two vertices being within hyperbolic distance R, given their radii. The lemma reduces a statement about hyperbolic distances to a statement about the relative angle between two points. Let us first introduce some notation. For a point $p \in \mathcal{D}_R$, we let $\theta(p) \in (-\pi, \pi]$ be the angle $p\hat{O}s$ between p and a (fixed) reference point $s \in \mathcal{D}_R$ (moving from s to p in the anti-clockwise direction). For $\theta, \theta' \in (-\pi, \pi]$, we set

$$|\theta - \theta'|_\pi = \min\{|\theta - \theta'|, 2\pi - |\theta - \theta'|\} \in [0, \pi].$$

For two points $p, p' \in \mathcal{D}_R$ we denote by $\theta(p, p') \in [0, \pi]$ their relative angle:

$$\theta(p, p') = |\theta(p) - \theta(p')|_\pi.$$

Also, for $p \in \mathcal{D}_R$ we let $y(p)$ denote the *defect radius* of p in $\mathcal{D}_R$. In other words, if $r(p)$ is the radius (the hyperbolic distance of p from O), then $y(p) = R - r(p)$. The following lemma gives a characterisation of what it is to have hyperbolic distance at most R in terms of the relative angle between two points. For r, r' such that $r + r' > R$, let $\theta_R(r, r') \in (-\pi, \pi]$ be such that if two points p, p' with $r(p) = r$ and $r(p') = r'$ have $\theta(p, p') = \theta_R(r, r')$ iff $d_H(p, p') = R$. In other words, $\theta_R(r, r')$ is the relative angle of two points of radii r and r', respectively, which are at distance R.

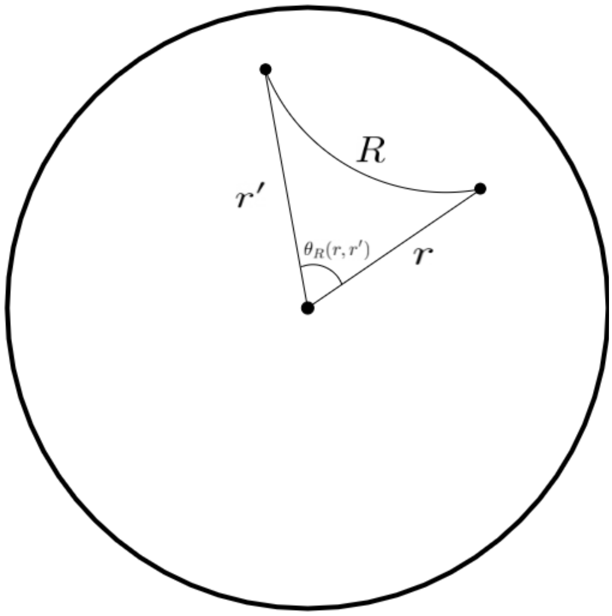

Figure 4.1: An example of $\theta_R(r, r')$ for two points at a distance of r and r' from the centre

Also, we set $T_R(y, y') = 2 \cdot e^{-R/2} e^{\frac{1}{2}(y+y')}$, for $y, y' \in [0, R]$. The following lemma is a consequence of Lemma 28 in [33].

Lemma 4.1.1. *Let $\kappa \in (0, 1)$. For any $\delta > 0$ and any n sufficiently large, uniformly for any $p, p' \in \mathcal{D}_R$ with $y(p) + y(p') \leq \kappa R$ the following holds*

$$\left| \frac{\theta_R(r(p), r(p'))}{T_R(y(p), y(p'))} - 1 \right| < \delta.$$

Proof. Lemma 28 in [33] states that there exists a constant $K > 0$ such that, for every $\varepsilon > 0$ and for R sufficiently large, the following holds: with

$$\Delta(r, r') := \frac{1}{2} e^{R/2} \arccos\left((\cosh r \cosh r' - \cosh R) / \sinh r \sinh r' \right),$$

For every $r, r' \in [\varepsilon R, R]$ with $r + r' > R$ we have that

$$e^{\frac{1}{2}(y+y')} - K e^{\frac{3}{2}(y+y') - R} \leq \Delta(r, r') \leq e^{\frac{1}{2}(y+y')} + K e^{\frac{3}{2}(y+y') - R}, \tag{4.1}$$

where $y := R - r, y' := R - r'$. The above is a consequence of the hyperbolic law of cosines.

Take two points p and p' inside $\mathcal{D}_R$ having radii r and r', respectively, and consider the triangle pOp'. If the distance between p and p' is R, then the angle opposite that side is $\theta_R(r,r')$ and satisfies:

$$\cosh(R) = \cosh r \cosh r' - \sinh r \sinh r' \cos(\theta_R(r,r')).$$

Therefore,

$$\theta_R(r,r') = 2e^{-R/2}\Delta(r,r').$$

So multiplying (4.1) by $2e^{-R/2}$ we get

$$T_R(y,y') - Ke^{\frac{3}{2}(y+y'-R)} \leq \theta_R(r,r') \leq T_R(y,y') + Ke^{\frac{3}{2}(y+y'-R)},$$

or

$$1 - \frac{K}{2}e^{y+y'-R} \leq \frac{\theta_R(r,r')}{T_R(y,y')} \leq 1 + \frac{K}{2}e^{y+y'-R}.$$

But $y+y' \leq \kappa R$ whereby $y+y'-R \leq (\kappa-1)R$. Since $\kappa < 1$, the lemma follows provided that n is sufficiently large. □

For a point $p \in \mathcal{D}_R$, we let $B(p;R)$ denote the set of points in $\mathcal{D}_R$ of hyperbolic distance at most R from p. We further define

$$\check{B}_{\kappa,\delta}(p) := \{p' \in \mathcal{D}_R \ : \ y(p') + y(p) \leq \kappa R, \theta(p,p') < (1+\delta)T_R(y(p), y(p'))\}.$$

We can think of $\check{B}_{\kappa,\delta}(p)$ to be an approximation to the ball $B(p;R)$. Essentially $B(p;R)$ is a subset of $\check{B}_{\kappa,\delta}(p)$, and in subsequent calculations $\check{B}_{\kappa,\delta}(p)$ is easier to work with. Let $\mathcal{A}_r := \mathcal{D}_R \setminus \mathcal{D}_r$ denote the annulus of the disc $\mathcal{D}_R$ which consists of all points of defect radius at most $R-r$. The above lemma implies that for any $\kappa \in (0,1)$, $\delta > 0$ and any n sufficiently large we have

$$\check{B}_{\kappa,-\delta}(p) \subset B(p;R) \cap \mathcal{A}_{(1-\kappa)R+y(p)} \subset \check{B}_{\kappa,\delta}(p); \tag{4.2}$$

hence, the set $\check{B}_{\kappa,\delta}(p)$ includes all points in $B(p;R)$ of defect radius at most $\kappa R - y(p)$. Furthermore, the following holds and will be useful later on during our second-moment calculations.

Claim 4.1.2. *If $\kappa \in (0,1)$ and $\delta > 0$, then for any n sufficiently large whenever $\theta(p,p') > 4(1+\delta)e^{-(1-\kappa)R/2}$ for points $p,p' \in \mathcal{D}_R$ with $y(p), y(p') < R/2$, we have*

$$\left(B(p;R) \cap \mathcal{A}_{(1-\kappa)R+y(p)}\right) \cap \left(B(p';R) \cap \mathcal{A}_{(1-\kappa)R+y(p')}\right) = \varnothing.$$

Proof. By (4.2), it is sufficient to show that

$$\check{B}_{\kappa,\delta}(p) \cap \check{B}_{\kappa,\delta}(p') = \varnothing.$$

Suppose not and let $p'' \in \check{B}_{\kappa,\delta}(p) \cap \check{B}_{\kappa,\delta}(p')$. Then by the definition of $\check{B}_{\kappa,\delta}(p)$ we have

$$\theta(p,p'') < 2(1+\delta)e^{R/2}e^{(y(p)+y(p''))/2} < 2(1+\delta)e^{-R/2+\kappa R/2}.$$

Similarly,

$$\theta(p',p'') < 2(1+\delta)e^{R/2}e^{(y(p')+y(p''))/2} < 2(1+\delta)e^{-R/2+\kappa R/2}.$$

So,

$$\theta(p,p') \leq \theta(p,p'') + \theta(p',p'') < 4(1+\delta)e^{-(1-\kappa)R/2}.$$

This concludes the proof of the claim. □

Another result, that will be useful later on, provided a bound on the expected number of points of $\mathrm{P}_{\alpha,\nu,n}$ inside $\check{B}_{\kappa,\delta}(p)$.

Claim 4.1.3. *For any $\kappa \in (1/2,1)$ and $\delta \in (-1,1)$, uniformly for any $p \in \mathcal{D}_R$ with $y(p) \leq R/2$ we have*

$$\mathbb{E}\left(|\mathrm{P}_{\alpha,\nu,n} \cap \check{B}_{\kappa,\delta}(p)|\right) = \Theta(e^{y(p)/2}).$$

Proof. We first observe that uniformly for all $0 < \rho < R$:

$$\frac{\alpha \sinh(\alpha\rho)}{\cosh(\alpha R) - 1} = \Theta\left(\frac{e^{\alpha\rho}}{e^{\alpha R} - 1}\right) = \Theta\left(e^{-\alpha(R-\rho)}\right).$$

Hence we may calculate the following

$$\begin{aligned}
&\mathbb{E}\left(|\mathrm{P}_{\alpha,\nu,n} \cap \check{B}_{\kappa,\delta}(p)|\right) = \\
&= n\frac{1+\delta}{2\pi} \cdot 2e^{-R/2+y(p)/2} \cdot \int_{(1-\kappa)R+y(p)}^{R} e^{(R-\varrho)/2} \frac{\alpha \sinh(\alpha\varrho)}{\cosh(\alpha R) - 1} d\varrho \\
&= \Theta(1) \cdot e^{y(p)/2} \int_{(1-\kappa)R+y(p)}^{R} e^{(1/2-\alpha)(R-\varrho)} d\varrho \\
&= \Theta(1) \cdot e^{y(p)/2} \int_{0}^{\kappa R - y(p)} e^{(1/2-\alpha)y} dy \overset{\alpha>1/2}{=} \Theta(e^{y(p)/2}).
\end{aligned}$$

□

Furthermore, $|\mathrm{P}_{\alpha,\nu,n} \cap \check{B}_{\kappa,\gamma}(p)|$ follows a Poisson distribution. For a random variable $X \sim \mathrm{Po}(\lambda)$ we have $\mathbb{E}(X^2) = \lambda^2 + \lambda$, hence the above claim also yields that for any $\kappa \in (1/2, 1)$ and $\delta > 0$,

$$\mathbb{E}\left(|\mathrm{P}_{\alpha,\nu,n} \cap \check{B}_{\kappa,\delta}(p)|^2\right) = O(e^{y(p)}), \tag{4.3}$$

uniformly for any $p \in \mathcal{D}_R$ with $y(p) \leq R/2$.

4.1.4 Preliminaries: A Projection of $\mathcal{D}_R$ onto $\mathbb{R}^2$

To simplify our calculations, we will transfer our analysis from $\mathcal{D}_R$ to $\mathbb{R}^2$. In particular, we make use of a mapping introduced in [33] which reduces our model to that of a percolation model on $\mathbb{R}^2$. Our result could in principle be proved without the use of it, but the proofs would be much heavier. This is achieved using a local approximation of the hyperbolic metric as given in Lemma 4.1.1. For a point, $p \in \mathcal{D}_R$, let $(\theta(p), y(p)) \in (-\pi, \pi] \times [0, R]$ denote its angle with respect to a reference point and its defect radius, respectively. We define the map $\Phi : \mathcal{D}_R \to \mathcal{B} = (-\frac{\pi}{2}e^{R/2}, \frac{\pi}{2}e^{R/2}] \times [0, R]$, mapping a point $p = (\theta(p), y(p)) \in$

$\mathcal{D}_R$ to a point $(x(p), y(p)) \in \mathcal{B}$

$$\theta(p) \mapsto x(p) := \frac{1}{2}\theta(p)e^{R/2} \text{ and } y(p) \mapsto y(p).$$

For simplicity, we set $I := I(R) := \frac{\pi}{2}e^{R/2}$. The map Φ projects the process $\mathrm{P}_{\alpha,\nu,n}$ to a point process on $\mathcal{B}$. We will approximate this process with the Poisson point process on $\mathcal{B}$ having intensity

$$\frac{2\nu}{\pi}\alpha e^{-\alpha y}dxdy.$$

For any measurable subset $S \subseteq \mathcal{B}$, we set $\mu_{\alpha,\beta}(S) = \beta \int_S e^{-\alpha y}dxdy$, with $\beta = \frac{2\nu\alpha}{\pi}$. We denote this Poisson process by $\mathrm{P}_{\alpha,\beta}$. The analogue of the relative angle between points in $\mathcal{D}_R$ is defined as follows. For $x, x' \in (-I, I]$, we let

$$|x - x'|_{\mathcal{B}} := \min\left\{|x - x'|, 2I - |x - x'|\right\}.$$

For a positive real number $y < R$, we set $\mathcal{B}(y) := (-\frac{\pi}{2}e^{R/2}, \frac{\pi}{2}e^{R/2}] \times [0, y]$; thus $\mathcal{B}(R) = \mathcal{B}$. We define the random graph $\mathcal{B}_y(n; \alpha, \nu)$ with vertex set the point set of $\mathrm{P}_{\alpha,\beta} \cap \mathcal{B}(y)$, and for any distinct $p, p' \in \mathrm{P}_{\alpha,\beta}$, the vertices p, p' are adjacent if and only if

$$|x(p) - x(p')|_{\mathcal{B}} < e^{(y(p)+y(p'))/2}.$$

We define the ball around a point $p \in \mathcal{B}(y)$ as $B_y(p) = \{p' \in \mathcal{B}(y) \ : \ |x(p) - x(p')|_{\mathcal{B}} < e^{\frac{1}{2}(y(p)+y(p'))}\}$. Thus, for a point $p \in \mathrm{P}_{\alpha,\beta}$, the neighbourhood of p in the random graph $\mathcal{B}_y(n; \alpha, \nu)$ is $B_y(p) \cap \mathrm{P}_{\alpha,\beta} \setminus \{p\}$. Figure 4.2 shows the neighbourhood around a point $p \in \mathcal{B}(y)$. Thus any point lying in the region bounded by the x-axis and the two log curves will be connected to p. The rectangular region bounded by the axis and the dotted line represents a single box in our partition, see Section 4.3.

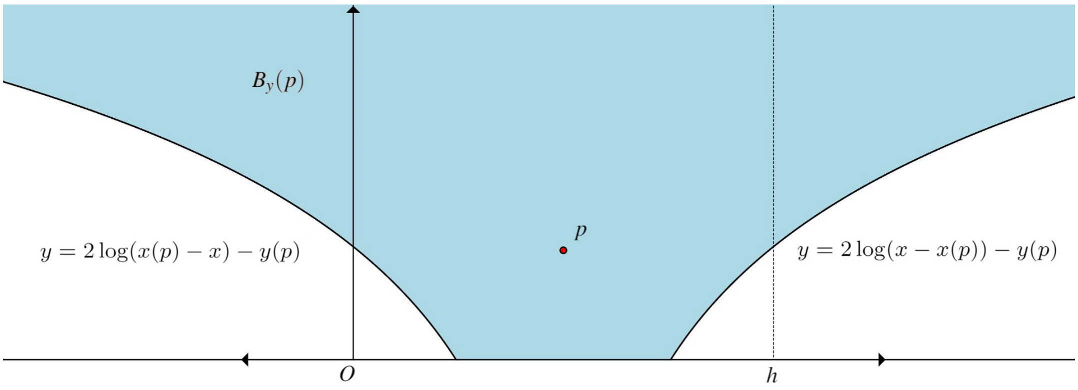

Figure 4.2: The ball $B_y(p)$.

4.2 Mapping $\mathcal{P}(n;\alpha,\nu)$ into $\mathcal{B}$ and the proof structure of Theorem 1.3.1

To prove Theorem 1.3.1, it suffices to consider a subgraph of $\mathcal{P}(n;\alpha,\nu)$ which contains most edges of it. To this end, we use Lemma 5.1 from [60].

Lemma 4.2.1. *Let $G = (V,E)$ be a graph with $|E| \geq 1$, let E_0 be a nonempty subset of E. For $E' = E \backslash E_0$, let $G' = (V,E')$. Then*

$$|\text{mod}(G) - \text{mod}(G')| < 2|E_0|/|E|.$$

We will show the following lemma.

Lemma 4.2.2. *Let $\alpha > 1/2$ and $\nu > 0$. For every $\varepsilon > 0$ there exists $y_\varepsilon > 0$ such that a.a.s.*

$$\text{vol}(\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{R-y_\varepsilon}) \leq \varepsilon e(\mathcal{P}(n;\alpha,\nu)).$$

For a positive real number $y < R$, let $\mathcal{P}_{\leq y}(n;\alpha,\nu)$ denote the subgraph of $\mathcal{P}(n;\alpha,\nu)$ induced by the points of $\mathrm{P}_{\alpha,\nu,n}$ having defect radius at most y. (The subgraph $\mathcal{P}_{>y}(n;\alpha,\nu)$ is defined analogously.) As the number of edges incident to points in $\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{R-y_\varepsilon}$ is at most $\text{vol}(\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{R-y_\varepsilon})$, the above two results imply that for every $\varepsilon > 0$ there exists

$y_\varepsilon > 0$ such that a.a.s.

$$|\mathrm{mod}(\mathcal{P}(n;\alpha,\nu)) - \mathrm{mod}(\mathcal{P}_{\le y_\varepsilon}(n;\alpha,\nu))| < 2\varepsilon.$$

Thereby, to prove Theorem 1.3.1 it suffices to show that

$$\mathrm{mod}(\mathcal{P}_{\le y_\varepsilon}(n;\alpha,\nu)) \to 1, \tag{4.4}$$

as $n \to \infty$ in probability. To show this, we will couple the random graph $\mathcal{P}_{\le y_\varepsilon}(n;\alpha,\nu)$ with the random graph $\mathcal{B}_{y_\varepsilon}(n;\alpha,\nu)$.

Lemma 4.2.3 Lemmas 27 and 30 in [33]**.** *There is a coupling between the point processes* $\mathrm{P}_{\alpha,\beta}$ *and* $\mathrm{P}_{\alpha,\nu,n}$ *such that a.a.s. on the coupling space* $\Phi(\mathrm{P}_{\alpha,\nu,n}) = \mathrm{P}_{\alpha,\beta}$*. Furthermore, a.a.s. on the coupling space for any distinct* $p, p' \in \mathrm{P}_{\alpha,\nu,n}$ *with* $y(p), y(p') \le R/4$ *we have* $d_H(p,p') \le R$ *if and only if* $\Phi(p') \in B(\Phi(p);R)$*.*

The above lemma implies that there is a coupling between the processes $\mathrm{P}_{\alpha,\beta}$ and $\Phi(\mathrm{P}_{\alpha,\nu,n})$ on $\mathcal{B}$ such that for any fixed $y > 0$ a.a.s., on this coupling space, the two point-sets coincide and moreover the random graph $\mathcal{B}_y(n;\alpha,\nu)$ is isomorphic to $\mathcal{P}_{\le y}(n;\alpha,\nu)$.

So we can deduce (4.4) from the following theorem.

Theorem 4.2.4. *For any* $\alpha > 1/2$*,* $\nu > 0$ *and any fixed* $y > 0$*, we have*

$$\mathrm{mod}(\mathcal{B}_y(n;\alpha,\nu)) \to 1,$$

as $n \to \infty$ *in probability.*

4.3 The Modularity of $\mathcal{B}_y(n;\alpha,\nu)$

In this section, we provide the proof of Theorem 4.2.4. Namely that the modularity of $\mathcal{B}_y(n;\alpha,\nu) \to 1$ as $n \to \infty$, in probability. We split the proof into two subsections.

The first subsection provides a lower bound on the modularity of G, given as a rough characterisation of how much G differs from an optimally modular graph. While in a later subsection, we calculate concentrations on edges within, and between, parts of a given partition of $\mathcal{B}_y(n;\alpha,\nu)$. Applying this, we obtain that for any $\varepsilon > 0$ it follows that for $\mathrm{mod}(\mathcal{B}_y(n;\alpha,\nu)) \geq 1-\varepsilon$ a.a.s.

4.3.1 Some General Properties of Graph Modularity

For a given graph G and $A, B \subset V$, let $\bar{A}$ denote $V \setminus A$ and, for disjoint A, B, let $e(A,B)$ denote the number of edges with one end-vertex in A and the other in B. We recall $e(A)$, which denotes the number of edges within part A, and $\mathrm{vol}(A) = \sum_{v\in A} \deg(v)$ which denotes the volume of A. Let $\mathcal{A}$ be a partition of V. We recall the modularity score of $\mathcal{A}$ in G is defined to be,

$$\mathrm{mod}_{\mathcal{A}}(G) = \sum_{A\in\mathcal{A}} \left(\frac{e(A)}{m} - \left(\frac{\mathrm{vol}(A)}{2m} \right)^2 \right)$$

It will sometimes be helpful to talk separately of the edge-contribution, also called *coverage*

$$\mathrm{mod}^E_{\mathcal{A}}(G) = \frac{1}{m} \sum_{A\in\mathcal{A}} e(A) = 1 - \frac{1}{2m} \sum_{A\in\mathcal{A}} e(A,\bar{A}),$$

and the *degree tax*

$$\mathrm{mod}^D_{\mathcal{A}}(G) = \frac{1}{(2m)^2} \sum_{A\in\mathcal{A}} \mathrm{vol}(A)^2.$$

The following lemma provides a lower bound on $\mathrm{mod}_{\mathcal{A}}(G)$ with respect to the parameters of a given partition $\mathcal{A}$.

Lemma 4.3.1. *Let G be a graph with m edges. Suppose the partition $\mathcal{A} = \{A_1, \ldots, A_k\}$ has the property that for each $1 \leq i \leq k$ and some $\varepsilon, \delta > 0$,*

$$e(A_i, \bar{A}_i) \leq \varepsilon m \quad \text{and} \quad |\mathrm{vol}(A_i) - 2m/k| \leq 2m\delta,$$

then

$$\text{mod}_{\mathcal{A}}(G) \geq 1 - \frac{k\varepsilon}{2} - \frac{1}{k} - k\delta^2.$$

Proof. Define δ_i to be such that $\text{vol}(A_i) = (1/k + \delta_i)2m$ and note that $\sum_i \delta_i = 0$ and by assumption $\forall i\ |\delta_i| \leq \delta$. We may now bound the degree tax of $\mathcal{A}$,

$$\text{mod}^D_{\mathcal{A}}(G) = \frac{1}{4m^2} \sum_{i=1}^{k} \text{vol}(A_i)^2 = \sum_{i=1}^{k} \left(\frac{1}{k} + \delta_i\right)^2 \leq \frac{1}{k} + k\delta^2.$$

The edge contribution of $\mathcal{A}$ is $\text{mod}^E_A(G) = 1 - \sum_i e(A_i, \bar{A}_i)/2m \geq 1 - k\varepsilon/2$ and thus we have our required bound. □

4.3.2 Proof of Theorem 4.2.4

We shall make use of the following identity which is an application of the (multivariate) *Campbell-Mecke* formula (see for example Theorem 4.4 [54]): for a Poisson point process $\mathcal{P}$ on a measurable space S with intensity ρ and a measurable non-negative function $h : S^k \times \mathcal{N} \to \mathbb{R}$, where $\mathcal{N}$ is the set of all locally finite collections of points in S, we have

$$\begin{aligned} &\mathbb{E}\left(\sum_{\substack{x_1,\ldots,x_k \in \mathcal{P} \\ \forall i,j, x_i \neq x_j}} h(x_1, \ldots, x_k, \mathcal{P} \setminus \{x_1, \ldots, x_k\})\right) \\ &= \int_S \cdots \int_S \mathbb{E}\left(h(x_1, \ldots, x_r, \mathcal{P} \cup \{x_1, \ldots, x_k\})\right) \rho(x_1) \cdots \rho(x_k) dx_1 \cdots dx_k. \end{aligned} \tag{4.5}$$

We set out to show that for any fixed $y > 0$, we have $\text{mod}(\mathcal{B}_y(n; \alpha, \nu)) \to 1$ in probability as $n \to \infty$. To this end, we will use Lemma 4.3.1 on a specific partition of the vertex set of $\mathcal{B}_y(n; \alpha, \nu)$. More specifically, we consider a partition of the box $\mathcal{B}_y = (-I, I] \times [0, y]$ into $2t$ boxes $B_i := (i \cdot hI, (i+1) \cdot hI] \times [0, y]$, for $i = -1/h, \ldots, 1/h - 1$, where $h = 1/t$ with $t \in \mathbb{N}$. Given this partition of the box $\mathcal{B}_y$, we let $A_i = \text{P}_{\alpha,\beta} \cap B_i$, for

$i = -t, \dots, t-1$. With $\mathcal{A} = \{A_{-t}, \dots, A_{t-1}\}$, we will show that a.a.s.

$$\text{mod}_{\mathcal{A}}(\mathcal{B}_y(n;\alpha,\nu)) \geq 1 - 4h - o(1). \tag{4.6}$$

Therefore, for $\varepsilon > 0$, we take $t \in \mathbb{N}$ to be such that $4h = 4/t < \varepsilon/2$ and deduce that a.a.s.

$$\text{mod}(\mathcal{B}_y(n;\alpha,\nu)) \geq 1 - \varepsilon.$$

Let us now proceed with the proof of (4.6). Firstly, note that since the random variables $\text{vol}(A_i)$ are identically distributed, with m denoting the number of edges of the random graph $\mathcal{B}_y(n;\alpha,\nu)$, we have

$$\mathbb{E}\left(\text{vol}(A_i)\right) = \frac{1}{2t}\mathbb{E}\left[\,\text{vol}(\mathrm{P}_{\alpha,\beta} \cap \mathcal{B}_y)\,\right] = \frac{\mathbb{E}\left(m\right)}{t}. \tag{4.7}$$

We will use a second moment argument to show that a.a.s. for each $i = -t, \dots, t-1$, we have

$$2m\left(\frac{h}{2} - 3h^2\right) \leq \text{vol}(A_i) \leq 2m\left(\frac{h}{2} + 3h^2\right). \tag{4.8}$$

Furthermore, we will show the following.

Claim 4.3.2. *There exists a constant C (depending on y) such that*

$$\mathbb{E}\left(e(A_1, \overline{A}_1)\right) < C.$$

By the union bound and Markov's inequality, this implies that a.a.s. for all $i = -t, \dots, t-1$

$$e(A_i, \overline{A}_i) < \log n.$$

Since a.a.s. $m = \Omega(n)$, we can then deduce (4.6) by applying Lemma 4.3.1 with $\varepsilon = \log^2 n/n$, $\delta = 3h^2$, and $k = t = 1/h$. We will deduce (4.8) from Chebyschev's inequality has shown that both the expectation and the variance of $\text{vol}(A_i)$ are of order n. These quantities also depend on t, though as t is a fixed constant we may omit this dependency and only concern ourselves with the asymptotic in n.

Claim 4.3.3. *We have*

$$\mathbb{E}\left(\text{vol}(A_1)\right) = \Theta(n) \text{ and } \text{Var}(\text{vol}(A_1)) = O(n).$$

Since the random variables $\text{vol}(A_i)$ are identically distributed, the first part of the above claim together with (4.7) imply that $\mathbb{E}(m) = \Theta(n)$ too. Furthermore, Chebyschev's inequality implies that a.a.s.

$$2\mathbb{E}(m)\left(\frac{h}{2} - h^2\right) \leq \text{vol}(A_1) \leq 2\mathbb{E}(m)\left(\frac{h}{2} + h^2\right).$$

In turn, the union bound implies that a.a.s. for all $i = -t, \ldots, t-1$, we have

$$2\mathbb{E}(m)\left(\frac{h}{2} - h^2\right) \leq \text{vol}(A_i) \leq 2\mathbb{E}(m)\left(\frac{h}{2} + h^2\right). \tag{4.9}$$

Furthermore, a.a.s $m \geq \mathbb{E}(m)(1 - h^3)$. Indeed, we have by Chebyshev's inequality that for each $i = -t, \ldots, t-1$

$$\mathbb{P}\left[|\text{vol}(A_i) - \mathbb{E}\left(\text{vol}(A_i)\right)| > h^3\mathbb{E}\left(\text{vol}(A_i)\right)\right] \leq \frac{\text{Var}(\text{vol}(A_i))}{h^6\mathbb{E}\left(\text{vol}(A_i)\right)^2} \overset{\textit{Claim } 4.3.3}{=} o(1).$$

Hence, by the union bound, we have that a.a.s $\text{vol}(A_i) \geq (1 - h^3)\mathbb{E}\left(\text{vol}(A_i)\right)$ for all $i = -t, \ldots, t-1$. Therefore by the Handshaking Lemma, $\sum_{i=-t}^{t-1} \text{vol}(A_i) = 2m$ whereby

$$2m = \sum_{-t \leq i \leq t-1} \text{vol}(A_i) \geq \sum_{-t \leq i \leq t-1} (1 - h^3)\mathbb{E}\left(\text{vol}(A_i)\right) = 2\mathbb{E}(m)(1 - h^3)$$

and,

$$2m = \sum_{-t \leq i \leq t-1} \text{vol}(A_i) \leq \sum_{-t \leq i \leq t-1} (1 + h^3)\mathbb{E}\left(\text{vol}(A_i)\right) = 2\mathbb{E}(m)(1 + h^3).$$

From the above, we deduce (4.8) since a.a.s. for all $i = -t, \ldots, t-1$

$$\mathrm{vol}(A_i) \le \left(\frac{h}{2} + h^2\right)(1-h^3)^{-1} m \le \left(\frac{h}{2} + h^2\right)(1+h^2)m \le \left(\frac{h}{2} + 3h^2\right) m,$$

provided that $t \ge 2$ (so that $h^2 > h^3 + h^5$ which is equivalent to $1 > h + h^3$ and holds if $h \le 1/2$), and

$$\mathrm{vol}(A_i) \ge \left(\frac{h}{2} - h^2\right)(1+h^3)^{-1} m \ge \left(\frac{h}{2} - h^2\right)(1-h^3)m \ge \left(\frac{h}{2} - 3h^2\right) m.$$

Proof of Claim 4.3.2. Firstly, let us point out that if a point $p \in A_1$ is far from the boundary of B_1, then it does not contribute to $e(A_1, \overline{A}_1)$. To quantify this, let us recall that for another $p' \in \mathcal{B}(y)$, if $|x(p') - x(p)|_{\mathcal{B}} > e^{\frac{1}{2}(y(p)+y)}$, then $p' \notin B_y(p)$. Since $y(p) \le y$ as well, we can further conclude that for any point $p' \in \mathcal{B}(y)$, if $|x(p') - x(p)|_{\mathcal{B}} > e^{y}$, then $p' \notin B_y(p)$. Hence, the only points $p \in A_1$ that may contribute to $e(A_1, \overline{A}_1)$ are such that $0 \le x(p) < e^y$ or $hI - e^y \le x(p) < hI$. Let $A_1^{(1)}$ denote the set of the former and $A_1^{(2)}$ the set of the latter. Hence,

$$\mathbb{E}\left(e(A_1, \overline{A}_1)\right) \le \mathbb{E}\left(\mathrm{vol}(A_1^{(1)})\right) + \mathbb{E}\left(\mathrm{vol}(A_1^{(2)})\right) = 2 \cdot \mathbb{E}\left(\mathrm{vol}(A_1^{(1)})\right),$$

where the last equality holds since the random variables $\mathrm{vol}(A_1^{(1)})$ and $\mathrm{vol}(A_1^{(2)})$ are identically distributed. For a finite set of points P and a point $p \in P$, we let $\deg(p; P) = |B_y(p) \cap P \setminus \{p\}|$. Now, we apply the Campbell-Mecke formula (4.5) and get

$$\begin{aligned}\mathbb{E}\left(\mathrm{vol}(A_1^{(1)})\right) &= \mathbb{E}\left(\sum_{p \in \mathrm{P}_{\alpha,\beta} \cap A_1^{(1)}} \deg(p; \mathrm{P}_{\alpha,\beta})\right) \\ &\stackrel{(4.5)}{=} \beta \cdot \int_0^y \int_0^{e^y} \mathbb{E}\left(\deg((x_0, y_0)); \mathrm{P}_{\alpha,\beta} \cup \{(x_0, y_0)\}\right) \cdot e^{-\alpha y_0} dx_0 dy_0.\end{aligned}$$

But $\mathbb{E}\left(\deg((x_0, y_0)); \mathrm{P}_{\alpha,\beta} \cup \{(x_0, y_0)\}\right) = |B_y((x_0, y_0)) \cap \mathrm{P}_{\alpha,\beta}| = O(1)$, uniformly over all

$x_0 \in (0, e^y]$ and $y_0 \in [0, y]$. So

$$\mathbb{E}\left(\mathrm{vol}(A_1^{(1)})\right) = O(1) \cdot \int_0^y \int_0^{e^y} e^{-\alpha y_0} dx_0 dy_0 = O(1).$$

□

Proof of Claim 4.3.3. We will calculate $\mathbb{E}(\mathrm{vol}(A_1))$ with the use of the Campbell-Mecke formula (4.5):

$$\begin{aligned}
\mathbb{E}(\mathrm{vol}(A_1)) &= \mathbb{E}\left(\sum_{p \in A_1 \cap \mathrm{P}_{\alpha,\beta}} \deg(p; \mathrm{P}_{\alpha,\beta} \cup \{p\})\right) \\
&= \beta \cdot \int_0^{hI} \int_0^y \mathbb{E}\left(\deg((x_0, y_0); \mathrm{P}_{\alpha,\beta} \cup \{(x_0, y_0)\})\right) e^{-\alpha y_0} dy_0 dx_0 \\
&= \beta h I \cdot \int_0^y \mathbb{E}\left(\deg((0, y_0); \mathrm{P}_{\alpha,\beta} \cup \{(0, y_0)\})\right) e^{-\alpha y_0} dy_0, \qquad (4.10)
\end{aligned}$$

since $\mathrm{P}_{\alpha,\beta}$ is homogeneous on the x-coordinate and $\deg((x_0, y_0); \mathrm{P}_{\alpha,\beta} \cup \{(x_0, y_0)\})$ are identically distributed with respect to x_0. Now,

$$\begin{aligned}
\mathbb{E}\left(\deg((0, y_0); \mathrm{P}_{\alpha,\beta} \cup \{(0, y_0)\})\right) &= 2\beta \cdot \int_0^y e^{(y_0 + y_0')/2} e^{-\alpha y_0'} dy_0' \\
&\overset{\alpha > 1/2}{=} \frac{2\beta}{\alpha - 1/2} e^{y_0/2} \left(1 - e^{-y(\alpha - 1/2)}\right).
\end{aligned} \qquad (4.11)$$

We substitute the integrand in (4.10) with the above expression and get

$$\begin{aligned}
\mathbb{E}(\mathrm{vol}(A_1)) &= hI \frac{2\beta^2}{\alpha - 1/2} \left(1 - e^{-y(\alpha - 1/2)}\right) \cdot \int_0^y e^{y_0/2 - \alpha y_0} dy_0 \\
&= 2hI \left[\frac{\beta}{\alpha - 1/2} \left(1 - e^{-y(\alpha - 1/2)}\right)\right]^2 = \Theta(n).
\end{aligned}$$

We calculate $\mathrm{Var}(\mathrm{vol}(A_1)) = \mathbb{E}(\mathrm{vol}(A_1)^2) - (\mathbb{E}(\mathrm{vol}(A_1)))^2$. For convenience, let $p_0 = (x_0, y_0)$ and similarly $p_0' = (x_0', y_0')$. With the use of the Campbell-Mecke formula (4.5).

We write

$$\mathbb{E}\left(\mathrm{vol}(A_1)^2\right) = \mathbb{E}\left(\sum_{p,p' \in \mathrm{P}_{\alpha,\beta} \cap B_1} \deg(p; \mathrm{P}_{\alpha,\beta}) \cdot \deg(p'; \mathrm{P}_{\alpha,\beta})\right) \overset{(4.5)}{=}$$

$$\int_0^y \int_0^{hI} \int_0^y \int_0^{hI} \mathbb{E}\left(\deg(p_0; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}) \times \deg(p_0'; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\})\right) \cdot e^{-\alpha y_0} e^{-\alpha y_0'} dx_0' dy_0' dx_0 dy_0. \tag{4.12}$$

We will now argue that for the majority of the pairs of points $p_0, p_0' \in B_1$, the expectation that is inside this integral factorises. Suppose without loss of generality that $x_0 < x_0'$. In this case, $B_y(p_0) \cap B_y(p_0') = \varnothing$ if and only if $x_0' - x_0 > e^{(y_0'+y)/2} + e^{(y_0+y)/2}$. So, if this is the case, the random variables $\deg(p_0; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\})$ and $\deg(p_0; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\})$ are independent. For given $y_0, y_0' \in [0, y]$, we let

$$S(y_0, y_0') = \{(x_0, x_0') \in (0, hI] \times (0, hI] \ : \ 0 < x_0' - x_0 \leq e^{(y_0'+y)/2} + e^{(y_0+y)/2}\}.$$

With this definition, we split the quadruple integral in (4.12) in the following way:

$$\begin{aligned} &\int_0^y \int_0^y \int_{(0,hI]\times(0,hI]\setminus S(y_0,y_0')} \mathbb{E}\left(\deg(p_0) \cdot \deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \times \\ &\qquad e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\ + &\int_0^y \int_0^y \int_{S(y_0,y_0')} \mathbb{E}\left(\deg(p_0) \cdot \deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \times \\ &\qquad e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0'. \end{aligned} \tag{4.13}$$

If $(x_0, x_0') \in (0, hI] \times (0, hI] \setminus S(y_0, y_0')$, then the random variables $\deg(p_0; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\})$ and $\deg(p_0'; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\})$ are independent. In the first integral, the integrand is

$$\begin{aligned} &\mathbb{E}\left(\deg(p_0) \cdot \deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \\ &= \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \cdot \mathbb{E}\left(\deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \\ &= \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta} \cup \{p_0\}\right) \cdot \mathbb{E}\left(\deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0'\}\right). \end{aligned}$$

Therefore, we can bound the first integral in (4.13) as follows:

$$\begin{aligned}
&\int_0^y \int_0^y \int_{(0,hI]\times(0,hI]\setminus S(y_0,y_0')} \mathbb{E}\left(\deg(p_0)\cdot\deg(p_0'); \mathrm{P}_{\alpha,\beta}\cup\{p_0,p_0'\}\right) \times \\
&\qquad e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
= &\int_0^y \int_0^y \int_{(0,hI]\times(0,hI]\setminus S(y_0,y_0')} \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta}\cup\{p_0\}\right) \times \\
&\qquad \mathbb{E}\left(\deg(p_0'); \mathrm{P}_{\alpha,\beta}\cup\{p_0'\}\right) e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
\le &\int_0^y \int_0^y \int_{(0,hI]\times(0,hI]} \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta}\cup\{p_0\}\right)\cdot \mathbb{E}\left(\deg(p_0'); \mathrm{P}_{\alpha,\beta}\cup\{p_0'\}\right) \times \\
&\qquad e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
= &\left(\int_0^y \int_0^{hI} \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta}\cup\{p_0\}\right) e^{-\alpha y_0} dx_0 dy_0\right)^2.
\end{aligned}$$

But by the Campbell-Mecke formula (4.5), the latter is

$$(\mathbb{E}\left(\mathrm{vol}(A_1)\right))^2 = \left(\int_0^y \int_0^{hI} \mathbb{E}\left(\deg(p_0); \mathrm{P}_{\alpha,\beta}\cup\{p_0\}\right) e^{-\alpha y_0} dx_0 dy_0\right)^2.$$

Now, let us consider the second integral in (4.13). We first consider the following claim:

Claim 4.3.4. *Let $y > 0$ be a constant. Then for every $y_0, y_0' \in [0, y]$ and $(x_0, x_0') \in S(y_0, y_0')$, we have*

$$\mathbb{E}\left(\deg(p_0)\cdot\deg(p_0'); \mathrm{P}_{\alpha,\beta}\cup\{p_0,p_0'\}\right) = O(1).$$

Proof. By recalling that $p_0 = (x_0, y_0)$ and $p_0' = (x_0', y_0')$, we note by the Cauchy-Schwarz inequality

$$\begin{aligned}
\mathbb{E}\left(\deg(p_0)\cdot\deg(p_0'); \mathrm{P}_{\alpha,\beta}\cup\{p_0,p_0'\}\right) \le \mathbb{E}\left(\deg(p_0)^2; \mathrm{P}_{\alpha,\beta}\cup\{p_0,p_0'\}\right)^{1/2} \times \\
\mathbb{E}\left(\deg(p_0')^2; \mathrm{P}_{\alpha,\beta}\cup\{p_0,p_0'\}\right)^{1/2}.
\end{aligned}$$

Furthermore, as conditioning on a single point in the process can only change the square of the degree by a constant factor, we have that

$$\mathbb{E}\left(\deg(p_0)^2; \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) = \Theta(\mathbb{E}\left(\deg(p_0)^2; \mathrm{P}_{\alpha,\beta} \cup \{p_0\}\right)).$$

As the intensity measure is homogeneous on the x co-ordinate, it suffices for us to show that for any $y_0 \in [0, y]$,

$$\mathbb{E}\left(\deg(0, y_0)^2; \mathrm{P}_{\alpha,\beta} \cup \{(0, y_0)\}\right) = O(1).$$

Indeed, we note that as a consequence of equation (4.11) and the definition of the point process, the random variable $\deg(0, y_0)$ on the probability space of the point process $\mathrm{P}_{\alpha,\beta} \cup \{(0, y_0)\}$ follows the Poisson distribution $\mathrm{Po}\left(\frac{2\beta}{\alpha-1/2}e^{y_0/2}\left(1 - e^{-y(\alpha-1/2)}\right)\right)$. Set λ to be the parameter of this Poisson distribution. Since y is fixed and $y_0 \leq y$ we have $\lambda = O(1)$. Using the second moment of the Poisson distribution, we have:

$$\mathbb{E}\left(\deg(0, y_0)^2; \mathrm{P}_{\alpha,\beta} \cup \{(0, y_0)\}\right) = \lambda^2 + \lambda = O(1).$$

□

We now return our attention to the second integral in (4.13),

$$\begin{aligned}
&\int_0^y \int_0^y \int_{S(y_0,y_0')} \mathbb{E}\left(\deg(p_0) \cdot \deg(p_0'); \mathrm{P}_{\alpha,\beta} \cup \{p_0, p_0'\}\right) \times \\
&\qquad e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
&= O(1) \cdot \int_0^y \int_0^y \int_{S(y_0,y_0')} e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
&= O(1) \cdot \int_0^y \int_0^y \int_0^{hI} \int_{x-2e^y}^{x+2e^y} e^{-\alpha y_0} e^{-\alpha y_0'} dx_0 dx_0' dy_0 dy_0' \\
&= O(1) \int_0^y \int_0^y \int_0^{hI} e^{-\alpha y_0} e^{-\alpha y_0'} dx_0' dy_0 dy_0' \\
&= O(n).
\end{aligned}$$

Thus, we conclude that

$$\mathbb{E}\left(\mathrm{vol}(A_1)^2\right) \leq \left(\mathbb{E}\left(\mathrm{vol}(A_1)\right)\right)^2 + O(n),$$

whereby

$$\mathrm{Var}(\mathrm{vol}(A_1)) = O(n).$$

□

4.4 Comparing the Disk to the Annulus: Proof of Lemma 4.2.2

Here, we return to the probability space associated with the random graph $\mathcal{P}(n;\alpha,\nu)$. In particular, we will work with a subset of the point process $\mathrm{P}_{\alpha,\nu,n}$ on $\mathcal{D}_R$, which we denote by $\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}$: we set $\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)} = \mathrm{P}_{\alpha,\nu,n} \setminus \mathcal{D}_{\delta R}$, for some $\delta \in (0,1)$. In other words, $\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}$ is $\mathrm{P}_{\alpha,\nu,n}$ but without the points inside the disc $\mathcal{D}_{\delta R}$. The reason for working with this process is that it is hard to bound the degrees of the points of $\mathrm{P}_{\alpha,\nu,n}$ which may appear close to the centre of $\mathcal{D}_R$. However, we can show that the two processes coincide a.a.s. provided that δ is small enough.

Claim 4.4.1. *If $\delta < 1 - 1/(2\alpha)$, then a.a.s.*

$$\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)} = \mathrm{P}_{\alpha,\nu,n}.$$

Proof. This follows from a simple first-moment argument. Indeed,

$$\mathbb{E}\left(|\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{\delta R}|\right) = n \cdot \kappa_{\alpha,\nu,n}(\mathcal{D}_{\delta R}) = n \cdot \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_0^{\delta R} \rho_n(r)\,dr\,d\theta.$$

But

$$\int_0^{\delta R} \rho_n(r)dr = \int_0^{\delta R} \frac{\alpha\sinh(\alpha r)}{\cosh(\alpha R)-1}dr = \frac{\cosh(\alpha\delta R)-1}{\cosh(\alpha R)-1} = O(n^{-2\alpha(1-\delta)}).$$

Therefore,

$$\mathbb{E}\left(|\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{\delta R}|\right) = O(n^{1-2\alpha(1-\delta)}).$$

So, if $\delta < 1 - 1/(2\alpha)$, then the exponent is negative and this expected value is $o(1)$. □

Note that $1 - 1/(2\alpha) < 1$, as $\alpha > 1/2$. Furthermore, note that the definition of $\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}$ allows for both processes to be defined on the same probability space, thus being naturally coupled. The intensity measure of $\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}$ is $n \cdot \kappa_{\alpha,\nu,n}(\cdot \setminus \mathcal{D}_{\delta R})$. For the moment, we shall assume that $\delta < 1 - 1/(2\alpha)$, so that the conclusion of Claim 4.4.1 holds.

For a point $p \in \mathcal{D}_R$ and a finite set of points $P \subset \mathcal{D}_R$, we set $\deg(p; P) = |B(p; R) \cap P \setminus \{p\}|$. For $0 \leq y_1 < y_2 \leq R$, we let $\mathcal{A}_{y_1,y_2} \subset \mathcal{D}_R$ denote the annulus inside $\mathcal{D}_R$ consisting of those points in $\mathcal{D}_R$ having defect radius between y_1 and y_2. We set

$$X_{y_1,y_2}(P) = \sum_{p \in P \cap \mathcal{A}_{y_1,y_2}} \deg(p; P).$$

Clearly, for any $0 < y < R$ on the event $\{\mathrm{P}_{\alpha,\nu,n} = \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\}$ we have

$$\mathrm{vol}(\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{R-y}) = X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})$$

and,

$$e(\mathcal{P}(n; \alpha, \nu)) = \frac{1}{2} X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}).$$

So, on $\{\mathrm{P}_{\alpha,\nu,n} = \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\}$, if $\mathrm{vol}(\mathrm{P}_{\alpha,\nu,n} \cap \mathcal{D}_{R-y}) > \varepsilon e(\mathcal{P}(n; \alpha, \nu))$, for some $\varepsilon > 0$, then

$$X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) > \frac{\varepsilon}{2} X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}). \tag{4.14}$$

We will give a general result on the concentration of the sum $X_{y,R}(\mathrm{P}_{\alpha,\nu,n})$, parametrised by y. We will show the following.

Lemma 4.4.2. *For any fixed $y \geq 0$, we have*

$$\frac{X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})}{\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)} \to 1,$$

as $n \to \infty$ in probability.

Furthermore, we show that $\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)$ decays exponentially in y.

Lemma 4.4.3. *For any $0 \leq y < R/4$ and any n sufficiently large, we have*

$$\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) \leq 2e^{-(\alpha-1/2)y} \cdot \mathbb{E}\left(X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right).$$

The above two lemmas imply that a.a.s.

$$X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) \leq 3e^{-(\alpha-1/2)y} X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}).$$

If we set $y = y_\varepsilon := \frac{1}{\alpha-1/2} \cdot \log(6/\varepsilon)$, it follows from (4.14) that

$$\begin{aligned} \mathbb{P}(e(\mathcal{P}_{>y_\varepsilon}(n;\alpha,\nu)) > \varepsilon e(\mathcal{P}(n;\alpha,\nu))) &\leq \mathbb{P}(X_{y_\varepsilon,R}(\mathrm{P}_{\alpha,\nu,n}) > \frac{\varepsilon}{2} X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})) \\ = \ \mathbb{P}(X_{y_\varepsilon,R}(\mathrm{P}_{\alpha,\nu,n}) > 3e^{-(\alpha-1/2)y_\varepsilon} X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})) = o(1). \end{aligned}$$

This concludes the proof of Lemma 4.2.2, assuming Lemmas 4.4.2 and 4.4.3. We now proceed with the proofs of these two lemmas.

Proof of Lemma 4.4.3. We lead with an upper bound on the expected value of the random variable $X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})$. Note that for $S < R$ we have $X_{y,S}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) \leq X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})$. So we can bound

$$\begin{aligned} 0 \ \leq \ X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) - X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) &\leq 2 \cdot |\{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{D}_{R/2}\}|^2 \\ &+ \sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{R/2,(1-\delta)R}} \deg(p; \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{0,R/2}). \end{aligned} \tag{4.15}$$

We will show that the right-hand side is sublinear a.a.s.

Claim 4.4.4. $\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) - \mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = o(n)$.

Proof. We lead by calculating the expected value of the first term on the right hand side of (4.15), given by $\mathbb{E}\left(|\{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{D}_{R/2}\}|^2\right)$. Thus we first calculate the first moment of $|\{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{D}_{R/2}\}|$ as,

$$\begin{aligned}
\mathbb{E}\left(|\{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{D}_{R/2}\}|\right) &= n \cdot \kappa_{\alpha,\nu,n}(\mathcal{D}_{R/2}) \\
&< n \cdot \frac{\alpha}{2\pi} \int_0^{R/2} \int_{-\pi}^{\pi} \frac{\sinh(\alpha r)}{\cosh(\alpha R) - 1} d\theta dr \\
&= n \cdot \frac{\cosh(\alpha R/2) - 1}{\cosh(\alpha R) - 1} = O(n^{1-\alpha}) \overset{\alpha > 1/2}{=} o(n^{1/2}).
\end{aligned}$$

Since this random variable is Poisson-distributed, the expected value of its square is proportional to the square of its expected value. Thereby,

$$\mathbb{E}\left(|\{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{D}_{R/2}\}|^2\right) = o(n). \tag{4.16}$$

We now bound the expected value of the last term in (4.15), using the Campbell-Mecke formula (4.5):

$$\mathbb{E}\left(\sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{R/2,(1-\delta)R}} \deg(p; \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{0,R/2})\right) =$$
$$n \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} \int_{\delta R}^{R/2} \mathbb{E}\left(\deg((\varrho, \theta); (\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho, \theta)\}) \cap \mathcal{A}_{0,R/2})\right) \rho_n(\varrho) d\varrho d\theta. \tag{4.17}$$

For a point $p = (\varrho, \theta) \in \mathcal{D}_R$ (here ϱ is the radius of p), we set $h_\kappa(p) := \kappa R - R + \varrho$. We will use the upper bound which is a consequence of (4.2): for $\kappa \in (0,1)$ and $\gamma \in (0,1)$

and for n sufficiently large

$$\deg((\varrho,\theta);(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\})\cap\mathcal{A}_{0,R/2})\leq|\check{B}_{\kappa,\gamma}((\varrho,\theta))\cap\mathcal{A}_{0,R/2}\cap\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}| \\ +|B^{\uparrow}((\varrho,\theta);\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})|,$$

where

$$B^{\uparrow}((\varrho,\theta);P):=\{p\in P\ :\ y(p)>h_{\kappa}((\varrho,\theta))\}.$$

Thereby,

$$\mathbb{E}\left(\deg((\varrho,\theta);(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\})\cap\mathcal{A}_{0,R/2})\right)\leq \\ \mathbb{E}\left(|\check{B}_{\kappa,\gamma}((\varrho,\theta))\cap\mathcal{A}_{0,R/2}\cap\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}|\right)+\mathbb{E}\left(|B^{\uparrow}((\varrho,\theta);\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})|\right). \tag{4.18}$$

Now, the first term on the right hand side of (4.18) can be bounded as follows:

$$\begin{aligned}\mathbb{E}\left(|\check{B}_{\kappa,\gamma}((\varrho,\theta))\cap\mathcal{A}_{0,R/2}\cap\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}|\right) &= \qquad (4.19)\\ &= n\frac{1+\gamma}{2\pi}\cdot 2e^{-R/2+(R-\varrho)/2}\cdot\int_{R/2}^{R}e^{(R-z)/2}\frac{\alpha\sinh(\alpha z)}{\cosh(\alpha R)-1}dz\\ &= \Theta(1)\cdot e^{(R-\varrho)/2}\int_{R/2}^{R}e^{(1/2-\alpha)(R-z)}dz\\ &= \Theta(1)\cdot e^{(R-\varrho)/2}\int_{0}^{R/2}e^{(1/2-\alpha)z}dz=\Theta(e^{(R-\varrho)/2}).\end{aligned}$$

Note that this bound is uniform over all $\varrho\in(\delta R,R/2)$. Substituting it in (4.17) we get

$$\begin{aligned}\mathbb{E}\left(\sum_{p\in\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cap\mathcal{A}_{R/2,(1-\delta)R}}|\check{B}_{\kappa,\gamma}((\varrho,\theta))\cap\mathcal{A}_{0,R/2}\cap\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}|\right) &=\\ &= O(n)\int_{-\pi}^{\pi}\int_{\delta R}^{R/2}e^{(R-\varrho)/2}\rho_n(\varrho)d\varrho d\theta\\ &= O(n)\int_{\delta R}^{R/2}e^{(R-\varrho)/2}e^{-\alpha(R-\varrho)}d\varrho\\ &\overset{\alpha>1/2}{=} O(n)e^{-(\alpha-1/2)R/2}=o(n).\end{aligned} \tag{4.20}$$

For the second term we have:

$$\begin{aligned}
&\mathbb{E}\left(|B^{\uparrow}((\varrho,\theta);\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})|\right) = n\cdot\kappa_{\alpha,\nu,n}(\{p\ :\ y(p)>\kappa R - R+\varrho\}) \\
&= \ n\cdot\frac{\alpha}{2\pi}\int_0^{2R-\kappa R-\varrho}\int_{-\pi}^{\pi}\frac{\sinh(\alpha r)}{\cosh(\alpha R)-1}d\theta dr \\
&= \ n\cdot\frac{\cosh(\alpha(2R-\kappa R-\varrho))-1}{\cosh(\alpha R)-1} = O(n\cdot e^{\alpha(R(1-\kappa)-\varrho)}),
\end{aligned} \tag{4.21}$$

uniformly over all $R/2<\varrho<R-y$. Therefore,

$$\begin{aligned}
n\cdot\frac{1}{2\pi}\int_{\delta R}^{R/2}\int_{-\pi}^{\pi}&\mathbb{E}\left(|B^{\uparrow}((\varrho,\theta);\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})|\right)\rho_n(\varrho)d\theta d\varrho = \\
&O(n^2)\cdot e^{\alpha R(1-\kappa)}\cdot\int_{\delta R}^{R/2}e^{-\alpha\varrho}\frac{\sinh(\alpha\varrho)}{\cosh(\alpha R)-1}d\varrho \\
\overset{\sinh(x)\le e^x}{\le}\quad &O(n^2)\cdot e^{\alpha R(1-\kappa)}\cdot\int_{\delta R}^{R/2}e^{-\alpha\varrho}\frac{e^{\alpha\varrho}}{\cosh(\alpha R)-1}d\varrho \\
= \quad &O(n^2)\cdot e^{\alpha R(1-\kappa)-\alpha R}\int_{\delta R}^{R/2}d\varrho \\
= \quad &O(R)\cdot n^2e^{-\alpha\kappa R} = O(R)\cdot n^{2(1-\alpha\kappa)}\overset{\alpha>1/2}{=}o(n),
\end{aligned}$$

provided that $1-\kappa$ is sufficiently small (depending on α). □

We can now consider $\mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)$. Applying the Campbell-Mecke identity (4.5) to the point process $\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}$ on $\mathcal{D}_R$ with intensity measure $n\cdot\kappa_{\alpha,\nu,n}(\cdot\setminus\mathcal{D}_{\delta R})$, we have

$$\begin{aligned}
\mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) &= \mathbb{E}\left(\sum_{p\in\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cap\mathcal{A}_{y,R/2}}\deg(p;\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = \\
n\cdot\frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{R/2}^{R-y}&\mathbb{E}\left(\deg((\varrho,\theta);(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\}))\right)\rho_n(\varrho)d\varrho d\theta.
\end{aligned} \tag{4.22}$$

Now, we bound the degree of (ϱ,θ) inside $\mathcal{A}_{y,R/2}$ with respect to the point process $\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\}$ with the use of Lemma 4.1.1. We apply (4.2) with κ sufficiently close to 1. For

$\gamma \in (0,1)$, and any finite subset $P \subset \mathcal{D}_R$ we bound

$$|\check{B}_{\kappa,-\gamma}((\varrho,\theta)) \cap P| \leq \deg((\varrho,\theta);P) \leq |B^{\uparrow}((\varrho,\theta);P)| + |\check{B}_{\kappa,\gamma}((\varrho,\theta)) \cap P|, \quad (4.23)$$

where

$$B^{\uparrow}((\varrho,\theta);P) := \{p \in P \ : \ y(p) > h_\kappa((\varrho,\theta))\}.$$

Let us set

$$X^{(\kappa,\gamma)}_{y_1,y_2}(P) = \sum_{p \in P \cap \mathcal{A}_{y_1,y_2}} |\check{B}_{\kappa,\gamma}(p) \cap P \setminus \{p\}|.$$

For the expected value of the first term in (4.23) we use the calculation in (4.21) which holds uniformly over all $R/2 < \varrho < R - y$:

$$n \cdot \frac{1}{2\pi} \int_{R/2}^{R-y} \int_{-\pi}^{\pi} \mathbb{E}\left(B^{\uparrow}((\varrho,\theta);(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\}) \cap \mathcal{D}_{R-y})\right) \rho_n(\varrho) d\theta d\varrho = o(n) \quad (4.24)$$

as in the proof of the previous claim, provided that $1-\kappa$ is sufficiently small (depending on α). Therefore,

$$0 \leq \mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) - \mathbb{E}\left(X^{(\kappa,\gamma)}_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = o(n). \quad (4.25)$$

Now, for any real γ such that $|\gamma| \in (0,1)$, we have

$$\mathbb{E}\left(|\check{B}_{\kappa,\gamma}((\varrho,\theta)) \cap \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}|\right) = n \cdot \frac{\alpha}{2\pi} \cdot (1+\gamma) 2e^{-R/2} \cdot e^{\frac{1}{2}(R-\varrho)} \int_{2R-\kappa R-\varrho}^{R} e^{\frac{1}{2}(R-z)} \frac{\sinh(\alpha z)}{\cosh(\alpha R)-1} dz.$$

For n sufficiently large we have

$$\left|\frac{\rho_n(z)}{e^{-\alpha(R-z)}} - 1\right| = \left|\frac{1}{e^{-\alpha(R-z)}} \cdot \frac{\alpha \sinh(\alpha z)}{\cosh(\alpha R)-1} - 1\right| < |\gamma|.$$

For real quantities a,b,c,d, with $c,d > 0$, we write that $a = d(b \pm c)$ if $d(b-c) \leq a \leq$

$d(b+c)$. So by the above inequality, the last integral is bounded, for n sufficiently large, as

$$\int_{2R-\kappa R-\varrho}^{R} e^{\frac{1}{2}(R-z)} \frac{\sinh(\alpha z)}{\cosh(\alpha R)-1} dz = \frac{(1 \pm |\gamma|)}{\alpha} \cdot \int_{2R-\kappa R-\varrho}^{R} e^{(\frac{1}{2}-\alpha)(R-z)} dz. \tag{4.26}$$

By applying the fact that $\varrho > R/2$ and $1-\kappa$ is sufficiently small (hence κ is bounded away from 1/2) we can compute the right hand integral as follows:

$$\begin{aligned} \int_{2R-\kappa R-\varrho}^{R} e^{(\frac{1}{2}-\alpha)(R-z)} dz &= \int_{0}^{\kappa R+\varrho-R} e^{(\frac{1}{2}-\alpha)z} dz \\ &\overset{\alpha>1/2}{=} \frac{1}{(\alpha-1/2)} \cdot \left(1-e^{(\frac{1}{2}-\alpha)(\kappa R+\varrho-R)}\right) \\ &= \frac{1}{(\alpha-1/2)}(1-o(1)). \end{aligned}$$

Therefore by substituting this expression into (4.26), and taking n to be sufficiently large for any $\varrho > R/2$

$$\int_{2R-\kappa R-\varrho}^{R} e^{\frac{1}{2}(R-z)} \frac{\sinh(\alpha z)}{\cosh(\alpha R)-1} dz = \frac{(1 \pm 2|\gamma|)}{\alpha(\alpha-1/2)}.$$

By substituting (4.26) and recalling that $\nu = ne^{-R/2}$, and setting $C_{\alpha,\nu} = \nu/(\pi(\alpha-1/2))$, it follows that uniformly for all $\varrho \in [R/2, R-y]$ and $\theta \in (-\pi, \pi]$ we have:

$$\frac{\mathbb{E}\left(|\breve{B}_{\kappa,\gamma}((\varrho,\theta)) \cap \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}|\right)}{e^{\frac{1}{2}(R-\varrho)}} = (1 \pm 2|\gamma|)^2 C_{\alpha,\nu}.$$

Therefore, by the Campbell-Mecke formula (4.5) we get:

$$\begin{aligned} \mathbb{E}\left(X_{y,R/2}^{(\kappa,\gamma)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right) &= \frac{n}{2\pi} \int_{R/2}^{R-y} \int_{-\pi}^{\pi} \mathbb{E}\left(|\breve{B}_{\kappa,\gamma}((\varrho,\theta)) \cap \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}|\right) \rho_n(\varrho) d\theta d\varrho \\ &= (1 \pm 2|\gamma|)^2 C_{\alpha,\nu} \cdot \frac{n}{2\pi} \int_{R/2}^{R-y} \int_{-\pi}^{\pi} e^{\frac{1}{2}(R-\varrho)-\alpha(R-\rho)} d\theta d\varrho. \end{aligned} \tag{4.27}$$

Again, we turn our attention to the right hand integral, as $\alpha > 1/2$ and $y < R/4$ we have the following:

$$\begin{aligned}
\int_{R/2}^{R-y}\int_{-\pi}^{\pi} e^{\frac{1}{2}(R-\varrho)-\alpha(R-\rho)}d\theta d\varrho &= 2\pi\int_{R/2}^{R-y} e^{(1/2-\alpha)(R-\varrho)}d\varrho \\
&= 2\pi\int_{y}^{R/2} e^{(1/2-\alpha)z}dz \\
&= \frac{2\pi}{\alpha-1/2}e^{-(\alpha-1/2)y}\left(1-e^{-(\alpha-1/2)(R/2-y)}\right) \\
&= \frac{2\pi}{\alpha-1/2}e^{-(\alpha-1/2)y}(1-o(1)),
\end{aligned}$$

uniformly over $y < R/4$. Substituting the above into (4.27), and taking n sufficiently large and setting $C'_{\alpha,\nu} = C_{\alpha,\nu}/(\alpha-1/2)$, we have the following:

$$\mathbb{E}\left(X^{(\kappa,\gamma)}_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = n(1\pm 3|\gamma|)^2 C'_{\alpha,\nu}e^{-(\alpha-1/2)y}.$$

So (4.22) and (4.25) yield, for sufficiently large n

$$\mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = n(1\pm 4|\gamma|)^2\cdot C'_{\alpha,\nu}e^{-(\alpha-1/2)y}.$$

Combining this with Claim 4.4.4 we deduce the following result: for $\gamma\in(-1,1)$, and n sufficiently large, we have for all $0\le y < R/4$,

$$n(1-5|\gamma|)^2C'_{\alpha,\nu}e^{-(\alpha-1/2)y} \le \mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) \le n(1+5|\gamma|)^2C'_{\alpha,\nu}e^{-(\alpha-1/2)y}. \tag{4.28}$$

By applying (4.28) we bound the following ratio: for $|\gamma|$ chosen small enough such that $(1+5|\gamma|)/(1-5|\gamma|) < \sqrt{2}$ and n sufficiently large: for all $0\le y < R/4$,

$$\begin{aligned}
\frac{\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)}{\mathbb{E}\left(X_{0,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)} &\le \frac{n(1+5|\gamma|)^2C'_{\alpha,\nu}e^{-(\alpha-1/2)y}}{n(1-5|\gamma|)^2C'_{\alpha,\nu}} = \frac{(1+5|\gamma|)^2}{(1-5|\gamma|)^2}e^{-(\alpha-1/2)y} \\
&\le 2e^{-(\alpha-1/2)y}.
\end{aligned}$$

$\square$

Proof of Lemma 4.4.2. Since

$$0 \leq \mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) - \mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = o(n),$$

but $\mathbb{E}\left(X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = \Theta(n)$ (for fixed $y > 0$) to show the concentration of $X_{y,R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})$ around its expected value, it suffices to show that

$$\frac{X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})}{\mathbb{E}\left(X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)} \to 1$$

as $n \to \infty$, in probability.

We decompose this random variable as follows:

$$X_{y,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) = X_{y,\log R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) + X_{\log R,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}).$$

By applying the upper bound of (4.28) with $y = \log R$, we deduce that

$$\mathbb{E}\left(X_{\log R,R/2}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right) = o(n).$$

For $p \in \mathcal{D}_R$, we set

$$\deg_{>h_\kappa(p)}(p;P) := |\{p' \in (P \setminus \{p\}) \cap B(p;R) \ : \ y(p') > h_\kappa(p)\}|$$

and

$$\deg_{\leq h_\kappa(p)}(p;P) := |\{p' \in (P \setminus \{p\}) \cap B(p;R) \ : \ y(p') \leq h_\kappa(p)\}|.$$

Hence, we express

$$X_{y,\log R}(\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}) \quad = \sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{y,\log R}} \deg_{>h_\kappa(p)}(p;\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})$$

$$+ \sum_{p \in \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)} \cap \mathcal{A}_{y,\log R}} \deg_{\leq h_\kappa(p)}(p; \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}).$$

Note that $\deg_{>h_\kappa(p)}(p; P) \leq |B^{\uparrow}(p; P)|$. So, by (4.24), the first term has

$$\mathbb{E}\left(\sum_{p \in \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)} \cap \mathcal{A}_{y,\log R}} \deg_{>h_\kappa(p)}(p; \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}) \right) = o(n).$$

For any $\kappa \in (0,1)$ and any finite set $P \subset \mathcal{D}_R$, set

$$X_{y,\log R}^{(\kappa)}(P) := \sum_{p \in P \cap \mathcal{A}_{y,\log R}} \deg_{\leq h_\kappa(p)}(p; P).$$

Therefore,

$$\mathbb{E}\left(X_{y,\log R}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right) = \mathbb{E}\left(X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right) + o(n).$$

In turn,

$$\mathbb{E}\left(X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right) = \Theta(n).$$

too. Hence, to show concentration around its expected value, it suffices to show that $X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})$ is concentrated around its expected value: as $n \to \infty$

$$\frac{X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})}{\mathbb{E}\left(X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right)} \to 1, \tag{4.29}$$

in probability. Since the expected value scales linearly in n, (4.29) will follow if we show that

$$\mathrm{Var}\left(X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})\right) = o(n^2).$$

4.4.1 Second-Moment Calculations

To bound the variance of $X_{y,\log R}^{(\kappa)}(\mathrm{P}_{\alpha,\nu,n}^{(>\delta R)})$, we will use Claim 4.1.2: we set $t_{\kappa,\gamma,R} := 4(1+\gamma)e^{-(1-\kappa)R/2}$ and write $\mathcal{A}_{y,\log R}^2$ for the product $\mathcal{A}_{y,\log R} \times \mathcal{A}_{y,\log R}$. We apply the Campbell-

Mecke formula (4.5)

$$\frac{(2\pi)^2}{n^2}\mathbb{E}\left(\left(\sum_{p\in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cap \mathcal{A}_{y,\log R}} \deg_{<h_\kappa(p)}(p;\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n})\right)^2\right) =$$

$$\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta))\cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'));\right.$$
$$\left.\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\}\right)\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho$$
$$=\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta))\cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'))\times\right.$$
$$\left.\mathbf{1}_{|\theta-\theta'|_\pi\le t_{\kappa,\gamma,R}};\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\}\right)\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho$$
$$+\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta))\cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'))\times\right.$$
$$\left.\mathbf{1}_{|\theta-\theta'|_\pi> t_{\kappa,\gamma,R}};\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\}\right)\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho.$$

Recall that for $r>0$, we defined $\mathcal{A}_r=\mathcal{A}_{0,R-r}$. To bound the second integral, let us observe that by Claim 4.1.2, if $|\theta-\theta'|_\pi > t_{\kappa,\gamma,R}$, then

$$\left(B_R((\varrho,\theta))\cap \mathcal{A}_{R-h_\kappa((\varrho,\theta))}\right)\cap\left(B_R((\varrho',\theta'))\cap \mathcal{A}_{R-h_\kappa((\varrho',\theta'))}\right)=\varnothing.$$

So, the random variable $\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta);\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\})$ and the random variable

$\deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta');\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\})$ are independent. Thus, we can write

$$\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta))\cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'))\times\right.$$
$$\left.\mathbf{1}_{|\theta-\theta'|_\pi> t_{\kappa,\gamma,R}};\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta),(\varrho',\theta')\}\right)\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho$$
$$=\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta));\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\}\right)\times$$
$$\mathbb{E}\left(\deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'));\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho',\theta')\}\right)$$
$$\mathbf{1}_{|\theta-\theta'|_\pi> t_{\kappa,\gamma,R}}\cdot\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho$$
$$\le\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta));\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho,\theta)\}\right)\times$$
$$\mathbb{E}\left(\deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta'));\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n}\cup\{(\varrho',\theta')\}\right)\rho_n(\varrho)\rho_n(\varrho')d\theta' d\varrho' d\theta d\varrho$$

$$= \left(\int_{\mathcal{A}_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\}\right) \cdot \rho_n(\varrho) d\varrho d\theta \right)^2$$

$$= \frac{(2\pi)^2}{n^2} \mathbb{E} \left(\sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{y,\log R}} \deg_{<h_\kappa(p)}(p) \right)^2,$$

by the Campbell-Mecke formula (4.5). For the first integral, we bound the product of the degrees by the sum of their squares:

$$\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)) \cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta')) \leq \deg^2_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)) + \deg^2_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta')).$$

So, by symmetry, we bound the first integral as follows:

$$\begin{aligned}
& \int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)) \cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta')) \times \right. \\
& \qquad \left. \mathbf{1}_{|\theta-\theta'|_\pi \leq t_{\kappa,\gamma,R}}; \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta),(\varrho',\theta')\}\right) \rho_n(\varrho)\rho_n(\varrho') d\theta' d\varrho' d\theta d\varrho \\
\leq\ & 2 \cdot \int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg^2_{<h_\kappa((\varrho,\theta))}((\varrho,\theta); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\})\right) \cdot \mathbf{1}_{|\theta-\theta'|_\pi \leq t_{\kappa,\gamma,R}} \times \\
& \qquad \rho_n(\varrho)\rho_n(\varrho') d\theta' d\varrho' d\theta d\varrho \\
=\ & 4 t_{\kappa,\gamma,R} \left(\int_{\mathcal{A}_{y,\log R}} \mathbb{E}\left(\deg^2_{<h_\kappa((\varrho,\theta))}((\varrho,\theta); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\})\right) \rho_n(\varrho) d\theta d\varrho \right) \times \\
& \qquad \left(\int_{\mathcal{A}_{y,\log R}} \rho_n(\varrho') d\theta' d\varrho' \right). \qquad (4.30)
\end{aligned}$$

But by (4.2), we have

$$\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\}) \leq |\mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \check{B}_{\kappa,\gamma}(p)|.$$

So by (4.3) we have

$$\mathbb{E}\left(\deg^2_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\}\right) = O(e^{R-\varrho}),$$

uniformly over all $R - \log R < \varrho < R - y$. Therefore,

$$\begin{aligned}
\int_{\mathcal{A}_{y,\log R}} \mathbb{E}\left(\deg^2_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)); \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta)\}\right) \rho_n(\varrho) d\theta d\varrho = \\
O(1) \cdot \int_{R-\log R}^{R-y} e^{R-\varrho} \frac{\sinh(\alpha\varrho)}{\cosh(\alpha R) - 1} d\varrho \\
= \quad O(1) \cdot \int_{R-\log R}^{R-y} e^{(R-\varrho)(1-\alpha)} d\varrho \\
= \quad O(1) \cdot \int_{y}^{\log R} e^{(1-\alpha)z} dz \overset{\alpha>1/2}{=} O(1) \cdot R^{1/2}. \qquad (4.31)
\end{aligned}$$

Furthermore,

$$\int_{\mathcal{A}_{y,\log R}} \rho_n(\varrho') d\theta' d\varrho' = 2\pi \frac{\cosh(\alpha(R-y)) - \cosh(\alpha(R - \log R))}{\cosh(\alpha R) - 1} = O(1). \qquad (4.32)$$

Using (4.31) and (4.32) into (4.30), we get

$$\begin{aligned}
\int_{\mathcal{A}^2_{y,\log R}} \mathbb{E}\left(\deg_{<h_\kappa((\varrho,\theta))}((\varrho,\theta)) \cdot \deg_{<h_\kappa((\varrho',\theta'))}((\varrho',\theta')) \times \right. \\
\left. \mathbf{1}_{|\theta-\theta'|_\pi \le t_{\kappa,\gamma,R}}; \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cup \{(\varrho,\theta),(\varrho',\theta')\}\right) \rho_n(\varrho)\rho_n(\varrho') d\theta' d\varrho' d\theta d\varrho \\
= \quad O(1) \cdot t_{\kappa,\gamma,R} R^{1/2} = O(1) \cdot e^{-(1-\kappa)R/2} R^{1/2} \\
= \quad O(1) \cdot n^{-(1-\kappa)} R^{1/2}.
\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}
\mathbb{E}\left(\left(\sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{y,\log R}} \deg_{<h_\kappa(p)}(p)\right)^2\right) \le \\
\mathbb{E}\left(\sum_{p \in \mathrm{P}^{(>\delta R)}_{\alpha,\nu,n} \cap \mathcal{A}_{y,\log R}} \deg_{<h_\kappa(p)}(p)\right)^2 + O(1) \cdot n^{2-(1-\kappa)} R^{1/2}.
\end{aligned}$$

Rearranging the above, we get

$$\mathrm{Var}\left(\sum_{p\in \mathrm{P}_{\alpha,\nu,n}^{(>\delta R)}\cap\mathcal{A}_{y,\log R}} \deg_{<h_\kappa(p)}(p)\right) = O(1)\cdot n^{1+\kappa}R^{1/2} = o(n^2).$$

□

4.5 Discussion

In this chapter, we have considered the modularity value of the KPKVB model of a random graph on the hyperbolic plane. We have shown that for all $\alpha > 1/2$ and $\nu > 0$ we have that $\mathrm{mod}(\mathcal{P}(n;\alpha,\nu)) \to 1$ as $n \to \infty$ in probability. The partition we consider is that of dividing the Poincaré disc into a constant number of equal sectors. We show that the modularity of this partition is closely related to the box partition given in $\mathcal{B}_y(n;\alpha,\nu)$. Following from this, we observe that for any $\varepsilon > 0$ a.a.s the modularity of $\mathcal{B}_y(n;\alpha,\nu)$ is at least $1-\varepsilon$ and thus $\mathrm{mod}(\mathcal{P}(n;\alpha,\nu)) \to 1$, as $n \to \infty$, in probability.

As we have now shown that the modularity value of the KPKVB model tends to 1 in probability, a natural question is the order of $1-\mathrm{mod}(G)$. This is also referred to as the *modularity deficit* of the graph [60]. The modularity deficit of a partition, $1-\mathrm{mod}_{\mathcal{A}}(G)$, quantifies how much a given partition $\mathcal{A}$ differs from optimal modularity. Trees on n nodes with maximum degree Δ have modularity deficit $O(\Delta^{\frac{1}{2}}n^{-\frac{1}{2}})$ [59], while the torus graph on n nodes in dimension d has deficit $O(n^{-1/2d})$ [62]. While we deduce that the modularity deficit of the sector division in the KPKVB model can be made arbitrarily small, it is open to determine whether we can explicitly express the rate of convergence asymptotically. It is also open to determine for a given growth rate, whether we can exhibit a partition that possesses such a deficit.

A modular community structure is characterised by a vertex partition where edge density within parts is much greater than expected, while density between parts is much smaller. While a high modularity value (> 0.3) can be indicative of an underlying modular community structure, a high value alone does not guarantee that such a community structure exists. This tends to occur in sparse networks. For example, in regimes where the average degree is bounded, the Erdős-Rényí random graph can exhibit a high modularity value in probability, without possessing a modular community structure [60].

In the case of the KPKBV model, the high modularity may be a consequence of the tree-like structure of the random graph. Generally, trees with sublinear maximum degree demonstrate an almost optimal modularity value; see [59]. Here, the term "tree-like" does not refer to the lack of short cycles (in fact, the presence of clustering implies that there are many short cycles with high probability). It refers to the existence of a hierarchy on the set of vertices of the random graph, which resembles the natural hierarchy that a rooted tree exhibits. Let us note that as a consequence of the negative curvature of hyperbolic space, tangential distances in the Poincaré disc expand exponentially with the respect to the radial distance from the centre. Pairs of vertices near the boundary of the disc are much less likely to connect, as they must possess an exponentially smaller relative angle for this to happen. In contrast, vertices near the centre have a relatively high degree, as the balls of radius R centred near O cover almost all of the disc. This implies that communities tend to have an underlying hierarchical structure, where they are formed from the mutual descendants of nodes with larger defect radii. Each part of the sector partition tends to capture a large proportion of one of these rooted sub-trees; therefore, this may suggest why the modularity score of the sector partition tends to one, in probability.

LIST OF REFERENCES

[1] M. Abdullah, M. Bode, and N. Fountoulakis. Typical distances in a geometric model for complex networks. *arXiv preprint arXiv:1506.07811*, 2015.

[2] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. A spatial web graph model with local influence regions. *Internet Mathematics*, 5(1-2):175–196, 2008.

[3] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.

[4] M. Alessandro and C. Vittorio. Leveraging the nonuniform pso network model as a benchmark for performance evaluation in community detection and link prediction. *New Journal of Physics*, 20(6):063022, 2018.

[5] N. Alon and J. Spencer. *The Probabilistic Method*. Wiley, third edition, 2008.

[6] A. Arenas, A. Fernandez, S. Fortunato, and S. Gomez. Motif-based communities in complex networks. *Journal of Physics A: Mathematical and Theoretical*, 41(22):224001, 2008.

[7] M. Bahrani, N. Immorlica, D. Mohan, and S.M. Weinberg. Asynchronous majority dynamics in preferential attachment trees. *arXiv preprint arXiv:1907.05823*, 33, 2019.

[8] A. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[9] I. Benjamini, S-O. Chan, R. O'Donnell, O. Tamuz, and L-Y. Tan. Convergence, unanimity and disagreement in majority dynamics on unimodular graphs and random graphs. *Stochastic Processes and their Applications*, 126:2719–2733, 2016.

[10] S. Blanda. The six degrees of kevin bacon. `https://blogs.ams.org/mathgradblog/2013/11/22/degrees-kevin-bacon/`, 2013. Date Accessed: 24/01/2023.

[11] V. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.

[12] L. E. Blume. The statistical mechanics of strategic interaction. *Games and Economic Behavior*, 5:387–424, 1993.

[13] M. Bode, N. Fountoulakis, and T. Müller. On the largest component of a hyperbolic model of complex networks. *The Electronic Journal of Combinatorics*, pages P3–24, 2015.

[14] M. Bode, N. Fountoulakis, and T. Müller. The probability of connectivity in a hyperbolic model of complex networks. *Random Structures & Algorithms*, 49(1):65–94, 2016.

[15] B. Bollobás. *Random graphs*. Cambridge University Press, second edition, 2001.

[16] U. Brandes, D. Delling, M. Gaertler, R. Görke, M. Hoefer, Z. Nikoloski, and D. Wagner. On finding graph clusterings with maximum modularity. In *International Workshop on Graph-Theoretic Concepts in Computer Science*, pages 121–132. Springer, 2007.

[17] P. Butkovič, D. Jones, and S. Sergeev. Lecture notes in game theory, February 2021.

[18] A. S. Chakrabarti, B. K. Chakrabarti, A. Chatterjee, and M. Mitra. The Kolkata paise restaurant problem and resource utilization. *Physica A: Statistical Mechanics and its Applications*, 388:2420–2426, 2009.

[19] D. Chakraborti, J. Han Kim, J. Lee, and T. Tran. Majority dynamics on sparse random graphs. *Random Structures & Algorithms*, 2021.

[20] D. Challet, A. Chessa, M. Marsili, and Y-C. Zhang. From minority games to real markets. *Quantitative Finance*, 1:168–176, 2001.

[21] D. Challet, M. Marsili, and Y-C. Zhang. *Minority games: interacting agents in financial markets*. Oxford University Press, 2005.

[22] J. Chellig, C. Durbac, and N. Fountoulakis. Best response dynamics on random graphs. *Games and Economic Behavior*, 131:141–170, 2022.

[23] J. Chellig, N. Fountoulakis, and F. Skerman. The modularity of random graphs on the hyperbolic plane. *Journal of Complex Networks*, 10(1):cnab051, 2022.

[24] A. Clauset, C. Moore, and M. Newman. Hierarchical structure and the prediction of missing links in networks. *Nature*, 453(7191):98–101, 2008.

[25] A. Clauset, M. Newman, and C. Moore. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004.

[26] L. Costa, F. Rodrigues, G. Travieso, and P. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances in physics*, 56(1):167–242, 2007.

[27] M.A.R. De Cara, O. Pla, and F. Guinea. Competition, efficiency and collective behavior in the "El Farol" bar model. *The European Physical Journal B-Condensed Matter and Complex Systems*, 10:187–191, 1999.

[28] T. Dinh, X. Li, and M. Thai. Network clustering via maximizing modularity: Approximation algorithms and theoretical limits. In *2015 IEEE International Conference on Data Mining*, pages 101–110. IEEE, 2015.

[29] P. Erdős and A. Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[30] P. Flajolet, D. Gardy, and L. Thimonier. Birthday paradox, coupon collectors, caching algorithms and self-organizing search. *Discrete Applied Mathematics*, 39(3):207–229, 1992.

[31] A. Fornito, A. Zalesky, and E. Bullmore. *Fundamentals of brain network analysis.* Academic press, 2016.

[32] N. Fountoulakis, M. Kang, and T. Makai. Resolution of a conjecture on majority dynamics: rapid stabilisation in dense random graphs. to appear in *Random Structures and Algorithms.*

[33] N. Fountoulakis and T. Müller. Law of large numbers for the largest component in a hyperbolic model of complex networks. *The Annals of Applied Probability*, 28(1):607–650, 2018.

[34] N. Fountoulakis, P. Van Der Hoorn, T. Müller, and M. Schepers. Clustering in a hyperbolic model of complex networks. *Electronic Journal of Probability*, 26:1–132, 2021.

[35] N. Fountoulakis and J. Yukich. Limit theory for isolated and extreme points in hyperbolic random geometric graphs. *Electronic Journal of Probability*, 25:1–51, 2020.

[36] T. Friedrich and A. Krohmer. On the diameter of hyperbolic random graphs. *SIAM Journal on Discrete Mathematics*, 32(2):1314–1334, 2018.

[37] E.N. Gilbert. Random graphs. *Annals of Mathematical Statistics*, 30:1141–1144, 1959.

[38] I. Gilboa and A. Matsui. Social stability and equilibrium. *Econometrica*, 59(3):859–867, 1991.

[39] E. Goles and J. Olivos. Periodic behaviour of generalized threshold functions. *Discrete Mathematics*, 30:187–189, 1980.

[40] M. Granovetter. The strength of weak ties. *American journal of sociology*, 78(6):1360–1380, 1973.

[41] L. Gugelmann, K. Panagiotou, and U. Peter. Random hyperbolic graphs: degree sequence and clustering. In *International Colloquium on Automata, Languages, and Programming*, pages 573–585. Springer, 2012.

[42] A. Hagberg, P. Swart, and D. Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[43] J. Hofbauer and K. Sigmund. Evolutionary game dynamics. *Bulletin of the American mathematical society*, 40:479–519, 2003.

[44] M. Hoffman, D. Steinley, K. Gates, M. Prinstein, and M. Brusco. Detecting clusters/communities in social networks. *Multivariate behavioral research*, 53(1):57–73, 2018.

[45] S. Janson, T. Łuczak, and A. Ruciński. *Random Graphs*. Wiley & Sons, first edition, 2000.

[46] W. Kets. The minority game: An economics perspective. *arXiv: 0706.4432*, 30, 2007.

[47] M. Kiwi and D. Mitsche. A bound for the diameter of random hyperbolic graphs. In *2015 Proceedings of the Twelfth Workshop on Analytic Algorithmics and Combinatorics (ANALCO)*, pages 26–39. SIAM, 2014.

[48] M. Kiwi and D. Mitsche. On the second largest component of random hyperbolic graphs. *SIAM Journal on Discrete Mathematics*, 33(4):2200–2217, 2019.

[49] B. Kovács and G. Palla. The inherent community structure of hyperbolic networks. *Scientific Reports*, 11(1):16050, 2021.

[50] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná. Hyperbolic geometry of complex networks. *Physical Review E*, 82(3):036106, 2010.

[51] M. Krivelevich. Long paths and hamiltonicity. `https://web.mat.bham.ac.uk/combinatorics/LMS-EPSRC/MichaelLectures3_4.pdf`, August 2013.

[52] M. Krivelevich and B. Sudakov. The phase transition in random graphs: A simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.

[53] M. Lasoń and M. Sulkowska. Modularity of minor-free graphs. *arXiv preprint arXiv:2102.07253*, 2021.

[54] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7. Cambridge University Press, 2017.

[55] M. Lelarge. Diffusion and cascading behavior in random networks. *ACM SIGMETRICS Performance Evaluation Review*, 39:34–36, 2011.

[56] L. Lichev and D. Mitsche. On the modularity of 3-regular random graphs and random graphs with given degree sequences. *Random Structures & Algorithms*, 2022.

[57] M. Marsili. Market mechanism and expectations in minority and majority games. *Physica A: Statistical Mechanics and its Applications*, 299:93–103, 2001.

[58] J. Maynard Smith and G.R. Price. The logic of animal conflict. *Nature*, 246:15–18, 1973.

[59] C. McDiarmid and F. Skerman. Modularity of regular and tree-like graphs. *Journal of Complex Networks*, 4:596–619., 2018.

[60] C. McDiarmid and F. Skerman. Modularity of erdős-rényi random graphs. *Random Structures & Algorithms*, 57(1):211–243, 2020.

[61] Stanley Milgram. The small world problem. *Psychology today*, 2(1):60–67, 1967.

[62] F. Montgolfier, M. Soto, and L. Viennot. Asymptotic modularity of some graph classes. In *International Symposium on Algorithms and Computation*, pages 435–444. Springer, 2011.

[63] G. Morrison, L. Dudte, and L. Mahadevan. Generalized erdős numbers for network analysis. *Royal Society Open Science*, 5(8):172281, 2018.

[64] E. Mossel, J. Neeman, and O. Tamuz. Majority dynamics and aggregation of information in social networks. *Autonomous Agents and Multi-Agent Systems*, 28(3):408–429, 2014.

[65] A. Mukhopadhyay, R. Mazumdar, and R. Roy. Voter and majority dynamics with biased and stubborn agents. *Journal of Statistical Physics*, 181(4):1239–1265, 2020.

[66] T. Müller and M. Staps. The diameter of kpkvb random graphs. *Advances in Applied Probability*, 51(2):358–377, 2019.

[67] A. Muscoloni and C. Cannistraci. A nonuniform popularity-similarity optimization (npso) model to efficiently generate realistic complex networks with communities. *New Journal of Physics*, 20(5):052002, 2018.

[68] R.B. Myerson. *Game theory.* Harvard university press, 2013.

[69] M. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, 74(3):036104, 2006.

[70] M. Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.

[71] M. Newman. *Networks*, volume 1. Oxford university press, 2018.

[72] M. Newman and M. Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.

[73] M. Nowak and R. M. May. Evolutionary games and spatial chaos. *Nature*, pages 826–829, 1992.

[74] M.J. Osborne et al. *An introduction to game theory*, volume 3. Oxford university press New York, 2004.

[75] F. Papadopoulos, M. Kitsak, M. Serrano, and D. Boguná, M.and Krioukov. Popularity versus similarity in growing networks. *Nature*, 489(7417):537–540, 2012.

[76] M. Penrose. *Random geometric graphs*, volume 5. OUP Oxford, 2003.

[77] V. V. Petrov. *Sums of independent random variables*, volume 82. Springer Science & Business Media, 2012.

[78] P. Pons and M. Latapy. Computing communities in large networks using random walks. In *International symposium on computer and information sciences*, pages 284–293. Springer, 2005.

[79] L. Pósa. Hamiltonian circuits in random graphs. *Discrete Mathematics*, 14(4):359–364, 1976.

[80] L. Prokhorenkova, P. Prałat, and A. Raigorodskii. Modularity in several random graph models. *Electronic Notes in Discrete Mathematics*, 61:947–953, 2017.

[81] U. Raghavan, R. Albert, and S. Kumara. Near linear time algorithm to detect community structures in large-scale networks. *Physical review E*, 76(3):036106, 2007.

[82] M. Rosvall and C. Bergstrom. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PloS one*, 6(4):e18209, 2011.

[83] F.C. Santos and J. M. Pacheco. Scale-free networks provide a unifying framework for the emergence of cooperation. *Phys. Rev. Lett.*, 95, 098104, 2005.

[84] C. Stegehuis, R. van Der Hofstad, and J. Van Leeuwaarden. Scale-free network clustering in hyperbolic and other random graphs. *Journal of Physics A: Mathematical and Theoretical*, 52(29):295101, 2019.

[85] J. von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton University Press, sixtieth-anniversary edition, 1944.

[86] X. Wang, G. Liu, J. Li, and J. Nees. Locating structural centers: A density-based clustering method for community detection. *PloS one*, 12(1):e0169355, 2017.

[87] E. Weisstein. Erfc. `https://mathworld.wolfram.com/Erfc.html`, 2023. Date Accessed: 27/09/2023.

[88] K. Zuev, M. Boguná, G. Bianconi, and D. Krioukov. Emergence of soft communities from geometric preferential attachment. *Scientific reports*, 5(1):1–9, 2015.